Europäisches Patentamt | European Patent Office | Office européen des brevets

Bescheinigung | Certificate | Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

03017677.0

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk

Anmeldung Nr:
Application no.: 03017677.0
Demande no:

Anmeldetag:
Date of filing: 14.08.03
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

Roche Vitamins AG

4070 Basel
SUISSE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

Microbial production of L-ascorbis acid

In Anspruch genommene Priorität(en) / Priority(ies) claimed /Priorité(s) revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

C12P/

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL
PT RO SE SI SK TR LI

Bemerkungen:
Remarks: The application was transferred from the above mentioned original applicant to:
Remarques: DSM IP Assets B.V - Heerlen, The Netherlands
The registration of the changes has taken effect on 28.11.2003.

Case 21864

MICROBIAL PRODUCTION OF L-ASCORBIC ACID

The present invention relates to polynucleotides derived from polynucleotides which encode ar enzyme which converts L-sorbosone directly to L-ascorbic acid. The enzyme L-sorbosone dehydrogenase (in the following: SNDHai) produces L-ascorbic acid (vitamin C) directly from L-sorbosone. The L-sorbosone dehydrogenase (SNDHai) was derived from bacteria belonging to the genera *Gluconobacter* and *Acetobacter*. L-Ascorbic acid is widely used in the pharmaceutical, food and cosmetic industries.

For the past 70 years, L-ascorbic acid (Vitamin C) has been produced industrially from D-glucose by the well-known Reichstein method. All steps in this process are chemical except for one (the conversion of D-sorbitol to L-sorbose) which is carried out by microbial transformation. Since its initial implementation for industrial production of L-ascorbic acid, several chemical and technical modifications have been used to improve the efficiency of the Reichstein method. Recent developments of vitamin C production are summarized in Ullmann's Encyclopedia of Industrial Chemistry, 5th Edition, Vol. A27 (1996), pp. 547ff. Recently different steps of vitamin C production have been performed with the help of microorganisms or enzymes isolated therefrom.

Current production methods for L-ascobic acid have some undesirable characteristics such as high energy consumption and use of large quantities of organic and inorganic solvents. Therefore, over the past decades, other approaches to manufacture L-ascorbic acid using microbial conversions, which would be more economical as well as ecological, have been investigated. Direct L-ascorbic acid production has been reported in several microorganisms.

Enzymatic conversion of L-sorbosone to L-ascorbic acid can be performed by an aldehyde dehydrogenase (SNDH) which is described in prior, non-prepublished patent applications (Cases 21422 and 21424). A further aldehyde dehydrogenase (SNDHII) that can convert L-sorbosone to L-ascorbic acid is described in another not prepublished patent application (Case

21246). Surprisingly, it has now been found that the direct conversion of L-sorbosone to L-ascorbic acid can be performed by acetic acid bacteria belonging to the genera *Gluconobacter* and *Acetobacter*. A gene responsible for this reaction was isolated and the sequence was determined. The sorbosone dehydrogenase enzyme encoded by this gene converts L-sorbosone to L-ascorbic acid and is hereafter called SNDHai. This enzyme is different from SNDH or SNDHII.

FIGURE 1 shows an alignment of the amino acid sequence of the *G. melanogenus* SNDHai and orthologous sequences.

As used herein, "cloning vector" means a plasmid or phage DNA or other DNA sequence which is able to replicate autonomously in a host cell, and which is characterized by one or a small number of restriction endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without loss of an essential biological function of the vector, and into which a DNA fragment may be spliced in order to bring about its replication. The cloning vector may further contain a marker suitable for use in the identification of cells transformed with the cloning vector. Markers, for example, may provide tetracycline resistance or ampicillin resistance.

As used herein, "expression" refers to the process by which a polypeptide is produced from a structural gene. The process involves transcription of the gene into mRNA and the translation of such mRNA into polypeptide(s).

As used herein, "expression vector" means a vector similar to a cloning vector but which is capable of enhancing the expression of a gene that has been cloned into it, after transformation into a host. The cloned gene is usually placed under the control of (i.e., operably linked to) certain control sequences such as promoter sequences. Promoter sequences may be either constitutive or inducible.

As used herein, "gene" refers to a DNA sequence that contains information needed for expressing a polypeptide or protein.

As used herein, "nucleic acid molecule" includes both DNA and RNA and, unless otherwise specified, includes both double-stranded, single-stranded nucleic acid, and nucleosides thereof. Also included are hybrids such as DNA-RNA hybrids, DNA-RNA-protein hybrids, RNA-protein hybrids, and DNA-protein hybrids. A polynucleotide consists of several bases, usually at least 20 nucleotide bases.

The term "homology" designates the similarity of two polynucleotide sequences. In order t determine the homology to polynucleotide sequences are arranged in such a manne that similar areas are compared. If required, nucleotides at certain positions may be replace by a blank position in order to improve the similarity. Homology comparisons can be performe by hand or by using computer programs which are commercially available. Preferably th program is run under standard conditions in order to obtain the maximum homology.

As used herein, "host" includes any prokaryotic or eukaryotic cell that is the recipient of a replicable expression vector or cloning vector. A "host," as the term is used herein, als includes a prokaryotic or eukaryotic cell that can be genetically engineered by well know techniques to contain desired gene(s) on its chromosome(s) or genome. Examples of suc hosts are known to the skilled artisan.

As used herein, "mutation" refers to a single base pair change, insertion or deletion in the nucleotide sequence of interest or to a genetic event such as an insertion of genetic elemen like a transposon.

As used herein, "mutagenesis" refers to a process whereby a mutation is generated in DNA. With "random" mutagenesis, the exact site of mutation is not predictable, occurring anywhere in the chromosome(s) of the microorganism or on the endogenous plasmid(s), and the mutation is brought about as a result of physical damage caused by agents such as radiation, chemical treatment, or insertion of a genetic element.

As used herein, "phenotype" refers to observable physical characteristics dependent upon the genetic constitution of a microorganism.

As used herein, "promoter" means DNA sequence generally described as the 5' region of a gene, located proximal to the start codon. The transcription of an adjacent gene(s) is initiated at the promoter region. If a promoter is an inducible promoter, then the rate of transcription increases in response to an inducing agent. In contrast, the rate of transcription is not regulated by an inducing agent if the promoter is a constitutive promoter.

As used herein, "recombinant" means a recombinant host which may be any prokaryotic or eukaryotic cell and contains the desired cloned gene(s) on an expression vector or cloning vector. This term also include those prokaryotic or eukaryotic cells that have been genetically engineered to contain the desired gene(s) in the chromosome or genome of that organism.

As used herein, "percent identical" refers to the percent of the amino acids of the subject amino acid sequence that have been matched to identical amino acids in the compared amino acid sequence by a sequence analysis program as exemplified below.

In one embodiment the present invention is related to an isolated polynucleotide derived from a polynucleotide molecule encoding a polypeptide having L-sorbosone dehydrogenase activity comprising a partial nucleotide sequence of SEQ ID No:1 of at least 20 consecutive nucleotides. The isolated polynucleotide comprises preferably a partial nucleotide sequence of at least 50 and more preferably of at least 100 consecutive nucleotides of SEQ ID No:1. The isolated polynucleotide is derived from a polynucleotide which codes for a polypeptide having L-sorbosone dehyrogenase activity. SEQ ID No:1 represents the complete nucleotide sequence of SNDHai which was isolated from the microorganism *Gluconobacter melanogenus*. Parts of such a sequence can be used for different purposes. Short polynucleotides for example can be used as primers for the amplification of suitable polynucleotides isolated from other organisms. Longer polynucleotides may code for polypeptides having enzymatic activity. For example, the SNDHai has a transmembrane domain which is not used for enzymatic activity. If parts of the polynucleotide coding for the enzymatically active area of the protein are expressed together with the transmembrane domain, such a polypeptide may have sufficient enzymatic activity.

The isolated polynucleotide molecule is usually derived from a longer polynucleotide sequence which codes also for a polypeptide having L-sorbosone dehydrogenase activity. Such polynucleotides may be isolated from bacteria preferably belonging to the genera *Gluconobacter* and *Acetobacter*. When such polynucleotides are derived from longer polynucleotide sequences it is possible to determine the homology between such a polynucleotide sequence and SEQ ID No:1. In such a case preferably an area having at least 100 consecutive nucleotides is selected and the corresponding stretch from the other polynucleotide is compared therewith. When the polynucleotide sequence and the corresponding stretch derived from SEQ ID No:1 have 60 nucleotides which are identical by comparing 100 consecutive nucleotides then the homology is 60%. The partial nucleotide sequences of the present invention are preferably derived from polynucleotide sequences having a homology of at least 80 and more preferably of at least 90% identity with SEQ ID No:1. For the determination of the homology stretches of at least 100, preferably at least 300 and more preferably at least 500 consecutive nucleotides are used.

The present invention provides novel polynucleotide sequences coding for L-sorbosone dehydrogenase of a microorganism belonging to acetic acid bacteria including the genera *Gluconobacter* and *Acetobacter*, for producing L-ascorbic acid from L-sorbosone. The said

polynucleotide preferably codes for a polypeptide having the amino acid sequence SEQ ID NO:2, or a polypeptide derived from that polypeptide by substitution, deletio insertion or addition of one or more amino acids in the amino acid sequence of SEQ ID NO: which retains L-sorbosone dehydrogenase activity to produce L-ascorbic acid from sorbosone.

The polypeptides of the present invention comprise preferably partial amino acid sequences at least 25 consecutive amino acids choosen from the amino acids sequences of th polypeptides disclosed in the present application. The person skilled in the art is aware of th fact that certain stretches in polypeptides are essential for the biologic activity. There are however, other areas wherein amino acids can be inserted, deleted or substituted by othe amino acids preferably such amino acids which are similar to the amino acids to be replaced.

This invention is further directed to recombinant DNA molecules and/or expression vector comprising a polynucleotide of the present invention, especially one which functions in a hos cell belonging to bacterial cells, yeast cells, fungal cells and plant cells. Preferably the host ce belongs to the genera *Gluconobacter*, *Acetobacter*, *Pseudomonas* or *Escherichia* or othe bacteria that can express the L-sorbosone dehydrogenase as an active form *in vivo*. This invention is also directed to recombinant organisms comprising such an expression vector o comprising such a nucleotide which has the polynucleotide integrated in its chromosomal DNA.

This invention is further directed to a process for producing recombinant L-sorbosone dehydrogenase polypeptide encoded by a polynucleotide of this invention by cultivating any of the recombinant organisms of this invention as described specifically above. Accordingly, part of this invention is a recombinant L-sorbosone dehydrogenase polypeptide produced by this process. In a preferred embodiment such recombinant L-sorbosone dehydrogenase is immobilized on a solid carrier for solid phase enzymatic reaction.

Another process of this invention is a process for producing L-ascorbic acid which comprises converting L-sorbosone into L-ascorbic acid with the aid of the recombinant L-sorbosone dehydrogenase polypeptide encoded by a polynucleotide of this invention.

In another embodiment of this invention a process for producing L-ascorbic acid comprises converting L-sorbose or D-sorbitol into L-ascorbic acid in a host cell having the ability for converting L-sorbose into L-sorbosone or for converting D-sorbitol into L-sorbosone, respectively.

The invention provides isolated nucleic acid molecules encoding the enzyme (L-sorbosone dehydrogenase SNDHai or parts thereof). Methods and techniques designed for the manipulation of isolated nucleic acid molecules are well known in the art. Methods for the isolation, purification, and cloning of nucleic acid molecules, as well as methods and techniques describing the use of eukaryotic and prokaryotic host cells and nucleic acid and protein expression therein, are known to the skilled person.

Functional derivatives of polypeptides of the present invention are defined on the basis of the amino acid sequences of the present invention by addition, insertion, deletion and/or substitution of one or more amino acid residues of such sequences wherein such derivatives preferably still have the L-sorbosone dehydrogenase activity measured by an assay known in the art or specifically described herein. Such functional derivatives can be made either by chemical peptide synthesis known in the art or by recombinant techniques on the basis of the DNA sequences as disclosed herein by methods known in the state of the art. Amino acid exchanges in proteins and peptides which do not generally alter the activity of such molecules are known.

In particular embodiments of the present invention, conservative substitutions of interest occur as follows: As example substitutions, Ala to Val/Leu/Ile, Arg to Lys/Gln/Asn, Asn to Gln/His/Lys/Arg, Asp to Glu, Cys to Ser, Gln to Asn, Glu to Asp, Gly to Pro/Ala, His to Asn/Gln/Lys/Arg, Ile to Leu/Val/Met/Ala/Phe/norLeu, Lys to Arg/Gln/Asn, Met to Leu/Phe/Ile, Phe to Leu/Val/Ile/Ala/Tyr, Pro to Ala, Ser to Thr, Thr to Ser, Trp to Tyr/Phe, Tyr to Trp/Phe/Thr/Ser, and Val to Ile/Leu/Met/Phe/Ala/norLeu are reasonable. As preferred examples, Ala to Val, Arg to Lys, Asn to Gln, Asp to Glu, Cys to Ser, Gln to Asn, Glu t Asp, Gly to Ala, His to Arg, Ile to Leu, Leu to Ile, Lys to Arg, Met to Leu, Phe to Leu, Pro to Ala, Ser to Thr, Thr to Ser, Trp to Tyr, Tyr to Phe, and Val to Leu are reasonable. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions described above, are introduced and the products screened.

Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequencer (such as the model Applied Biosystems PRISM 310 genetic analyzer). Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods

well known in the art. As is also known in the art, a single insertion or deletion in determined nucleotide sequence compared to the actual sequence will cause a frame shift translation of the nucleotide sequence such that the predicted amino acid sequence encod by a determined nucleotide sequence will be completely different from the amino ac sequence actually encoded by the sequenced DNA molecule, beginning at the point of such a insertion or deletion.

In a preferred embodiment the present invention is directed to polynucleotides encodin polypeptide having the L-sorbosone dehydrogenase activity as disclosed in the sequence listin as SEQ ID NO:2 as well as the complementary strands, or those which include thes sequences, DNA sequences or fragments thereof, and DNA sequences, which hybridize unde standard conditions with such sequences but which encode for polypeptides having exactly th same amino acid sequence.

Another mode of describing the similarity of polynucleotide sequences is to determine whethe such sequences do hybridize or do not hybridize. This depends from the conditions selected fo the hybridization.

Standard conditions for hybridization mean in this context such conditions which are generally used by a person skilled in the art to detect specific hybridization signals, or preferably so called "stringent hybridization conditions" used by a person skilled in the art. Thus, as used herein, the term "stringent hybridization conditions" means that hybridization will occur if there is about 95% and preferably at least 97% identity between the sequences. Stringent hybridization conditions are, e.g., 2 hrs to 4 days incubation at 42°C using a digoxigenin (DIG)-labeled DNA probe (prepared by using a DIG labeling system; Roche Dignostics GmbH, 68298 Mannheim, Germany) in a solution such as DigEasyHyb solution (Roche Diagnostics) with or without 100 µg/ml salmon sperm DNA, or a solution comprising 50% formamide, 5 × SSC (150 mM NaCl, 15 mM trisodium citrate), 0.02% sodium dodecyl sulfate, 0.1% N-lauroylsarcosine, and 2% blocking reagent (Roche Diagnostics GmbH), followed by washing the filters twice for 5 to 15 minutes in 2 x SSC and 0.1% SDS at room temperature and then washing twice for 15-30 minutes in 0.5 × SSC and 0.1% SDS or 0.1 x SSC and 0.1% SDS at 65-68°C.

Briefly, the gene encoding L-sorbosone dehydrogenase, the nucleic acid molecule containing said gene, the expression vector and the recombinant organism used in the present invention was obtained by the following steps:

(1) Perform transposon mutagenesis as described below on the strains belonging to the genus *Gluconobacter or Acetobacter* strains that produces L-ascorbic acid from L-sorbosone to obtain colonies expressing anitibiotic resistance encoded by the transposon used
(2) Select L-ascorbic acid non producing mutants in the screening with L-sorbosone as a substrate
(3) Isolate chromosomal DNA from the mutants
(4) Clone the DNA fragment containing the transposon from the chromosomal DNA by colony-, plaque-, or Southern-hybridization, PCR (polymerase chain reaction) cloning, and so on
(5) Determine the nucleotide sequence of the DNA fragment containing the transposon insertion
(6) Clone the DNA fragment from the parent strain that produces L-ascorbic acid from L-sorbosone
(7) Construct the expression vector on which the gene coding for L-sorbosone dehydrogenase can express efficiently
(8) Construct recombinant organisms carrying the gene coding for L-sorbosone dehydrogenase by an appropriate method for introducing DNA into host cell, e.g. transformation, transduction, conjugal transfer and/or electroporation, which host cell thereby becomes a recombinant organism of this invention.

The above-mentioned transposon mutagenesis is described in more detail below.

Transposon mutagenesis is known as a potent tool for genetic analysis (P. Gerhardt et al., "Methods for General and Molecular Bacteriology" Chapter 17, Transposon Mutagenesis; American Society for Microbiology). This method utilizes transposable elements which are distinct DNA segments having the unique capacity to move (transpose) to new sites within the genome of the host organisms. The transposition process is independent of the classical homologous recombination system of the organism. The insertion of a transposable element into a new genomic site does not require extensive DNA homology between the ends of the element and its target site. Transposable elements have been found in a wide variety of prokaryotic and eukaryotic organisms, where they can cause null mutations, chromosome rearrangements, and novel patterns of gene expression on insertion in the coding region or regulatory sequences of resident genes and operons.

Prokaryotic transposable elements can be roughly divided into three different classes. Class I consists of simple elements such as insertion sequences (IS elements), which are approximately 800 to 1,500 bp in length. IS elements normally consist of a gene encoding an enzyme required for transposition (i.e. transposase), flanked by terminally repeated DNA

sequences which serve as substrate for the transposase. IS elements were initia identified in the lactose and galactose utilization operons of enteric bacteria, whe the elements were found to cause often unstable, polar mutation on insertion.

Class II consists of composite transposable elements. The members of this class are als referred to as transposons or Tn elements. Transposons in prokaryotes have been identified a a class of complex transposable elements, often containing simple IS elements (or part thereof) as direct or inverted repeats at their termini, behaving formally like IS elements bu carrying additional genes unrelated to transposition functions, such as antibiotic resistance heavy-metal resistance or pathogenicity determinant genes. The insertion of a transposon int a particular genetic locus or replicon (phage) is designated by using a double colon, e.g *lacZ*::Tn5 or P1::Tn5.

Class III includes "transposable" bacteriophages, such as Mu and its relatives. Phage Mu is both a virus and a transposon. It is known that it can integrate at multiple sites in the host chromosome, thereby frequently causing mutations.

The transposon mutagenesis utilizing the above transposable elements is known to possess the following characteristics:

(i) Such mutation generally leads to inactivation of the gene, and the resulting null mutation is relatively stable.

(ii) Transposons introduce new genetic and physical markers into the target locus, such as antibiotic resistance genes, new restriction endonuclease cleavage sites, and unique DNA sequences which can be identified by genetic means, DNA-DNA hybridization, or electron microscopic heteroduplex analysis. The genetic markers are useful for mapping the mutated loci as well as screening the mutants.

(iii) Transposons can generate a variety of genomic rearrangements, such as deletions, inversions, translocations, or duplications, and can be used to introduce specific genes into the target bacteria.

A variety of transposons are known in the art, such as Tn3, Tn5, Tn7, Tn9, Tn10, phage Mu and the like. Among them, Tn5 is known to have almost no insertion specificity, and its size is relatively small. Tn5 is also one of the most frequently used transposable elements which is readily derived from the sources, such as pfd-Tn5 [American Type Culture Collection, USA (ATCC) ATCC 77330] or pCHR81 (ATCC 37535). For the purpose of use in the random mutagenesis in the practice of the present invention, Tn5 is preferred. A variety of Tn5 derivatives, designated Mini-Tn5s, which consists of 19 bp of the Tn5 inverted repeats required for transposition coupled to antibiotic resistance or other selectable marker genes are also useful for the present invention. Such Mini-Tn5s are inserted into a suicide vector, in addition to the Tn5 transposase (tnp), to construct an efficient suicide Tn5 mutagenesis system. Further information how to work with Tn5 transposons can be taken from the following references, P. Gerhardt et al., "Methods for General and Molecular Bacteriology" Chapter 17, Transposon Mutagenesis; American Society for Microbiology, 1994; K.N. Timmis et al., Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in Gram-negative Eubacteria, J. Bacteriology, 172: 6568-6572, 1990.

Information describing how Tn5 can be derived from pfd-Tn5 see explanation of pfd-Tn5 (ATCC 77330) obtainable from the ATCC Home Page on Internet, [http://www.atcc.org/catalogs/recomb.html]. Accordingly the suicide plasmid pfd-Tn5 can be introduced into *E. coli* as well as other Gram negative bacterium by electroporation (see the reference for recommended conditions). The plasmid itself can be used as a Tn5 donor. In addition, one can construct a suicide Tn5 vector by introducing pfd-Tn5 plasmid into the recipient *E. coli* with any plasmid which one wants to use, selecting transformants showing Km and resistance of the target plasmid, isolating plasmids from the transformants, transforming *E. coli* with the isolated plasmids, and selecting for Km and plasmid-marker-resistant transformants to obtain *E. coli* strain carrying the target palsmid with Tn5. Concept of this protocol is also available in "region-directed Tn5 mutagenesis" in P. Gerhardt et al. (mentioned above).

Random mutagenesis with transposon involves the introduction of a transposon into a target bacterial cell via transformation, transduction, conjugal mating or electroporation by using suicide plasmid or phage vectors. The resulting mutants may be screened with the aid of the marker carried by the transposon. Transposition of the transposon into the genome of the recipient bacterium can be detected after the vector used has been lost by segregation.

For the introduction of transposons into a microorganisms of the genus *Gluconobac* or *Acetobacter*, so-called suicide vectors including a derivative of phage P1 a narrow-host-range plasmids such as a derivative of pBR325 carrying ColE1 replication ori are commonly used. The phage P1 vectors and the plasmid vectors can be transferred infection and by transformation, conjugal mating or electroporation, respectively, into t recipient cells, wherein these vectors preferably lack the appropriate origins of recipients. T choice of suicide vector and transposon to be used depends on criteria including pha sensitivity, intrinsic antibiotic resistance of the recipient cell, the availability of a gene transf system including transformation, conjugal transfer, electroporation, or infection to introdu transposon-carrying vector into *E. coli*.

One of the preferable vectors for use in the present invention is phage P1 (ATCC25404) whic injects its DNA into a microorganism belonging to the genus *Gluconobacter* or *Acetobacte* however, this DNA will be unable to replicate and will be lost by segregation. Such P1 phag carrying Tn5 (P1::Tn5) can be used in the form of phage lysate which may be prepared b

lysing *E. coli* carrying P1::Tn5 in accordance with known procedures (see: e.g. Methods fo General and Molecular Bacteriology" Chapter 17, Transposon Mutagenesis; American Societ for Microbiology or US patent 5082785, 1992).

The other preferable suicide vectors which can be used in the present invention are plasmi suicide vectors, carrying the broad-host-range conjugal transfer of mobilization functions an sites but a narrow-host-range origin of replication. These vectors can be mobilized at a hig frequency from *E. coli* to a microorganism belonging to the genus *Gluconobacter* o *Acetobacter* but cannot be stably maintained in the recipient cells. These vectors contain, i addition to Tn5, the IncP-type mobilization (*mob*; *oriT*) site and are based on commonly used *E. coli* cloning vectors, such as pACYC177 (ATCC37031), pACYC184 (ATCC37033), and pBR32 [Bolivar F., 1978, Gene 4: 121-136; a derivative of pBR322 (ATCC31344)], all of which cannot replicate in nonenteric bacteria (pSUPseries, Simon R. et al., 1983, Bio/Technology 1: 784-791). pSUP-type plasmids such as pSUP2021 (Simon R. et al.) can be mobilized in bi-parental mating experiments by providing the transfer function in trans from a chromosomally integrated copy of the IncP plasmid RP4 in the donor strain itself (e. g., strain S17-1) or in tri-parental mating experiments by providing the transfer functions from plasmid pRK2013 (ATCC37159) harbored by a nondonor, nonrecipient helper strain of *E. coli*.

The recipient cell which received the transposon can be selected by the marker carried by the element, e.g. resistance to particular antibiotics. When Tn5 is used as a transposon, the marker of Km or Nm can be usually used. Besides Km or Nm markers, genetic markers of Tc, Gm, Sp, Ap, Cm and the like can be used as the alternative marker genes. The other transposons carrying readily visualized gene products such as those encoded by *lacZ*, *luxAB* or *phoA* can also be used. These Tn5 derivatives are especially useful if the target bacterial strain has an intrinsic resistance to the antibiotics normally used to select for Tn5 (kanamycin, neomycin, bleomycin and streptomycin) or if a secondary mutagenesis of a strain already harboring a Tn5 derivative is to be carried out (P. Gerhardt et al., "Methods for General and Molecular Bacteriology" Chapter 17, Transposon Mutagenesis; American Society for Microbiology).

For confirming that the deficient mutant really carries transposon, colony- or Southern-hybridization is usually conducted with labeled-DNA fragment containing the transposon used as the probe by the standard methods (Molecular cloning, a laboratory manual second edition, Maniatis T., et al., 1989).

Such a mutant was isolated as described in Example 3 of the present invention. The transposon mutant is useful for further identifying the target gene coding for L-sorbosone dehydrogenase and determining the nucleotide sequence of the region tagged with the transposon.

The DNA fragment containing a transposon insertion can be cloned into any *E. coli* cloning vectors, preferably pUC18, pUC19, pBluescript II KS+ (Stratagene Europe) and their relatives, by selecting transformants showing both phenotypes of selection markers of the vector and the transposon. The nucleotide sequences adjacent to the transposon are able to be determined by e.g., a chain termination method (Sanger F. S., et al., Proc. Natl. Acad. Sci. USA 74:5463-5467,1977) by using a kit such as Applied Biosystems ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kits (version 2).

Alternatively, when the said L-sorbosone dehydrogenase polypeptide is purified from the strain producing L-ascorbic acid from L-sorbosone, the desired gene can be cloned in either plasmid or phage vectors from a total chromosomal DNA by the following illustrative methods:
(i) The partial amino acid sequences can be determined from the purified proteins or peptide fragments thereof. Such whole protein or peptide fragments can be prepared by the isolation of such a whole protein or by peptidase-treatment from the gel after SDS-polyacrylamide gel electrophoresis. Thus obtained protein or fragments thereof are applied to protein sequencer

such as Applied Biosystems automatic gas- phase sequencer 470A. The amino ac sequences can be utilized to design and prepare oligonucleotide probes and/or prime with DNA synthesizer such as Applied Biosystems automatic DNA sequencer 381A. Tl probes can be used for isolating clones carrying the target gene from a gene library of tl strain carrying the target gene by means of Southern-, colony- or plaque-hybridization.
(ii) Further alternatively, for the purpose of selecting clones expressing target protein from th gene library, immunological methods with antibody prepared against the target protein can b applied.
(iii) The DNA fragment of the target gene can be amplified from the total chromosomal DNA b PCR method with a set of primers, i.e. two oligonucleotides synthesized according to the amin acid sequences determined as above. Then a clone carrying the target-whole gene can b isolated from the gene library constructed, e.g. in *E. coli* by Southern-, colony-, or plaque hybridization with the PCR product obtained above as the probe.

DNA sequences which can be made by PCR by using primers designed on the basis of th DNA sequences disclosed therein by methods known in the art are also an object of th present invention.

Above-mentioned antibody can be prepared with the purified L-sorbosone dehydrogenase proteins, the purified recombinant L-sorbosone dehydrogenase proteins such as His-tagged L-sorbosone dehydrogenase expressed in *E. coli*, or its peptide fragment as an antigen. Polypeptide sequence deduced from a nucleotide sequence of the L-sorbosone dehydrogenase can be used as an antigen for preparation of an antibody.
Once a clone carrying the desired gene is obtained, the nucleotide sequence of the target gene can be determined by a well known method.

To express the desired gene/nucleotide sequence efficiently, various promoters can be used; e.g., the original promoter of the gene, promoters of antibiotic resistance genes such as kanamycin resistant gene of Tn5, ampicillin resistant gene of pBR322, and beta-galactosidase of *E. coli* (*lac*), *trp*-, *tac*-, *trc*-promoter, promoters of lambda phage and any promoters which can be functional in a host cell. For this purpose, the host cell can be selected from a group consisting of bacterial cells, plant cells, fungal cells and yeast cells. Preferably the host cell belongs to the genera *Gluconobacter*, *Acetobacter*, *Pseudomonas*, *Escherichia* or bacteria that can express the L-sorbosone dehydrogenase as an active form *in vivo*, and most preferably it belongs to *Gluconobacter and Acetobacter.*

For expression, other regulatory elements, such as a Shine-Dalgarno (SD) sequence (e.g., AGGAGG and so on including natural and synthetic sequences operable in the host cell) and a transcriptional terminator (inverted repeat structure including any natural and synthetic sequence operable in the host cell) which are operable in the host cell (into which the coding sequence will be introduced to provide a recombinant cell of this invention) can be used with the above described promoters.

A wide variety of host/cloning vector combinations may be employed in cloning the double stranded DNA. Preferred vectors for the expression of the gene of the present invention in *E. coli* is selected from any vectors usually used in *E. coli*, such as pQE vectors which can express His-tagged recombinant proteins (QIAGEN AG Switzerland), pBR322 or its derivatives including pUC18 and pBluescript II (Stratagene Cloning Systems, Calif., USA), pACYC177 and pACYC184 and their derivatives, and a vector derived from a broad host range plasmid such as RK2 and RSF1010. A preferred vector for the expression of the nucleotide sequence of the present invention in bacteria including *Gluconobacter*, *Acetobacter*, *Pseudomonas* is selected from any vectors which can replicate in *Gluconobacter*, *Acetobacter*, *Pseudomonas* as well as in a preferred cloning organism such as *E. coli*. The preferred vector is a broad-host-range vector such as a cosmid vector like pVK100 and its derivatives and RSF1010. Copy number and stability of the vector should be carefully considered for stable and efficient expression of the cloned gene and also for efficient cultivation of the host cell carrying the cloned gene. Nucleic acid molecules containing transposable elements such as Tn5 can also be used as a vector to introduce the desired gene into the preferred host, especially on a chromosome. Nucleic acid molecules containing any DNAs isolated from the preferred host together with the gene of the present invention are also useful to introduce this gene into the preferred host cell, especially on a chromosome. Such nucleic acid molecules can be transferred to the preferred host by applying any of a conventional method, e.g., transformation, transduction, conjugal mating or electroporation, which are well known in the art, considering the nature of the host cell and the nucleic acid molecule.

The L-sorbosone dehydrogenase gene/nucleotide sequences provided in this invention are ligated into a suitable vector containing a regulatory region such as a promoter, a ribosomal binding site, and a transcriptional terminator operable in the host cell described above with a well-known method in the art to produce an expression vector.

To construct a recombinant microorganism carrying an expression vector, various gene transfer methods including transformation, transduction, conjugal mating, and electroporation can be used. The method for constructing a recombinant cell may be selected from the

methods well-known in the field of molecular biology. Conventional transformation syste can be used for *Gluconobacter*, *Acetobacter*, *Pseudomonas*, or *Escherichia*. A transduc system can also be used for *E. coli*. Conjugal mating system can be widely used in Gra positive and Gram-negative bacteria including *E. coli*, *P. putida*, and *Gluconobacter*. example of conjugal mating is disclosed in WO 89/06,688. The conjugation can occur in liq medium or on a solid surface. Examples for a recipient for SNDHai production incl microorganisms of *Gluconobacter*, *Acetobacter*, *Pseudomonas*, or *Escherichia*: To recipient for conjugal mating, a selective marker may be added; e.g., resistance agai nalidixic acid or rifampicin is usually selected. Natural resistance can also be used; e resistance against polymyxin B is useful for many *Gluconobacters*.

The present invention provides recombinant L-sorbosone dehydrogenase (SNDHai). One c increase the production yield of the L-sorbosone dehydrogenase enzyme by introducing the sorbosone dehydrogenase gene provided by the present invention into host cells descrit above. One can also produce more efficiently the L-sorbosone dehydrogenase proteins ir host cell selected from a group consisting of *Gluconobacter*, *Acetobacter*, *Pseudomonas*, *Escherichia* by using the L-sorbosone dehydrogenase gene of the present invention.

The microorganism may be cultured in an aqueous medium supplemented with appropri nutrients under aerobic conditions. The cultivation may be conducted at a pH of 4.0 to 9 preferably 5.0 to 8.0. The cultivation period varies depending on the pH, temperature a nutrient medium to be used, and is preferably about 1 to 5 days in batch or fed-batch mo The cultivation can be conducted in continuous mode with growing- or resting-cells. T preferred temperature range for carrying out the cultivation is from about 13°C to about 36 preferably from about 18°C to about 33°C. It is usually required that the culture medi contains such nutrients as assimilable carbon sources, e.g., glycerol, D-mannitol, D-sorbi erythritol, ribitol, xylitol, arabitol, inositol, dulcitol, D-ribose, D-fructose, D-glucose, and sucro preferably D-sorbitol, D-mannitol, and glycerol; and digestible nitrogen sources such as orga substances, e.g., peptone, yeast extract, baker's yeast, urea, amino acids, and corn ste liquor. Various inorganic substances may also be used as nitrogen sources, e.g., nitrates a ammonium salts. Furthermore, the culture medium usually contains inorganic salts, e. magnesium sulfate, potassium phosphate, and calcium carbonate.

The present invention provides production of L-ascorbic acid. One can produce L-ascorbic a from L-sorbosone, L-sorbose, or D-sorbitol by cultivating a recombinant organism provided the present invention described above. One can also produce L-ascorbic acid from sorbosone, L-sorbose, or D-sorbitol more efficiently by cultivating a recombinant organism w

a host cell selected from a group consisting of *Gluconobacter*, *Acetobacter*, *Pseudomonas* or *Escherichia* carrying L-sorbosone dehydrogenase gene of the present invention The recombinant microorganism may be cultured in an aqueous medium supplemented with appropriate nutrients under aerobic conditions. The cultivation may be conducted at a pH of 4.0 to 9.0, preferably 5.0 to 8.0. The cultivation period varies depending on the pH, temperature and nutrient medium to be used, and is preferably about 1 to 10 days in batch or fed-batch mode. The cultivation can be conducted in continuous mode with growting- or resting-cells. The preferred temperature range for carrying out the cultivation is from about 13°C to about 36°C preferably from about 18°C to about 33°C.

It is usually required that the culture medium contains such nutrients as assimilable carbon sources, e.g., glycerol, D-mannitol, D-sorbitol, erythritol, ribitol, xylitol, arabitol, inositol, dulcitol D-ribose, D-fructose, D-glucose, and sucrose, preferably D-sorbitol, D-mannitol, and glycerol and digestible nitrogen sources such as organic substances, e.g., peptone, yeast extract baker's yeast, urea, amino acids, and corn steep liquor. Various inorganic substances may also be used as nitrogen sources, e.g., nitrates and ammonium salts. Furthermore, the culture medium usually contains inorganic salts, e.g., magnesium sulfate, potassium phosphate, and calcium carbonate. The recombinant microorganism may be incubated under a resting cell condition, e.g., in a reaction mixture containing 0.3% NaCl, 1% CaCO3, and 0.1% to 50% of L-sorbosone, L-sorbose, or D-sorbitol. The reaction, which is also called as biotransformation may be conducted at a pH of 4.0 to 9.0, preferably 5.0 to 8.0.

The reaction period varies depending on the pH, temperature and reaction mixture to be used and is preferably about 1 to 10 days in batch or fed-batch mode. The reaction can be conducted in continuous mode. The preferred temperature range for carrying out the reaction is from about 13°C to about 36°C, preferably from about 18°C to about 33°C.

An embodiment for the isolation and purification of the recombinant L-sorbosone dehydrogenase from the microorganism after the cultivation is briefly described as follows. Cells are harvested from the liquid culture broth by centrifugation or filtration. The harvested cells are washed with water, physiological saline of a buffer solution having an appropriate pH. The washed cells are suspended in the buffer solution and disrupted by means of a homogenizer, sonicator or French press, or by treatment with lysozyme and so on to give a solution or disrupted cells. The recombinant L-sorbosone dehydrogenase was isolated and purified from the cell-free extract or disrupted cells by usual protein purification methods such as ultracentrifugation, ammonium sulfate precipitation, dialysis, ion exchange chromatographies, hydrophobic chromatographies, gel filtration chromatographies, affinity

chromatographies, and crystallization. When the recombinant L-sorbosone dehydrogenas is produced as tagged polypeptide such as His-tag one, it is purified with affinity res such as Ni affinity resin.

The recombinant L-sorbosone dehydrogenase as well as the recombinant organism can b immobilized on an appropriate solid carrier for solid phase reaction. Any means of immobilizin enzymes or organisms generally known in the art may be used. For instance, the enzyme ma be bound directly to a membrane, granules or the like of a resin having one or more functiona groups, or it may be bound to the resin through bridging compounds having one or mor functional groups, for example, glutaraldehyde.

The produced L-ascorbic acid in the reaction mixture or in the culture may be isolated b conventional methods known in the art, and it may be separated as a salt, e.g. of sodium potassium, calcium, ammonium or the like. This salt may be converted into a free acid b conventional methods known in the art. Specifically, the separation may be performed by an suitable combination or repetition of the following steps: formation of a salt by using differences in properties between the product and the surrounding impurities, such as solubility absorbability and distribution coefficient between the solvents, and absorption, for example, or ion exchange resin and the like. Any of these procedures alone or in combination constitutes a convenient means for isolating the product. The product thus obtained may further be purified in a conventional manner, e.g. by re-crystallization or chromatography.

The following Examples further illustrate the present invention and are not intended to limit the invention in any way.

Example 1. Production of L-ascorbic acid from L-sorbosone

Gluconobacter melanogenus IFO 3293 and *Gluconobacter melanogenus* N44-1, a derivative of the strain IFO 3293, (Sugisawa T. et al., Agric. Biol. Chem. 54: 1201-1209, 1990), *Gluconobacter melanogenus* IFO 3292, *Gluconobacter rubiginosus* IFO 3244, *Gluconobacter industrius* IFO 3260, *Gluconobacter cerinus* IFO 3266, *Gluconobacter oxydans* IFO 3287, *Acetobacter aceti subsp. orleanus* IFO 3259, *Acetobacter aceti subsp. xylinum* IFO 13693, *Acetobacter aceti subsp. xylinum* IFO 13773, and *Acetobacter sp.* ATCC 15164 were used for the production of L-ascorbic acid from L-sorbosone.

A. aceti subsp. xylinum IFO 13693 and *A. aceti subsp. xylinum* IFO 13773 were grown at 27°C for 3 days on No. 350 medium containing 5 g/l Bactopeptone (Difco), 5 g/L yeast extract (Difco), 5 g/L glucose, 5 g/L mannitol, 1 g/L $MgSO_4 \cdot 7H_2O$, 5 ml/L ethanol, and 15 g/L agar. All other *Acetobacter* strains and all *Gluconobacter* strains were grown at 27°C for 3 days on mannitol broth (MB) agar medium containing 25 g/l mannitol, 5 g/l yeast extract (Difco Laboratories, Detroit, Mich., USA), 3 g/l Bactopeptone (Difco), and 18 g/l of agar (Difco).

Cells were scraped from the agar plates, suspended in distilled water and used for resting cell reactions conducted at 30°C for 20 h in 5 ml tubes with shaking at 230 rpm. The reaction mixtures (0.5 ml) contained 1% L-sorbosone, 0.3% NaCl, 1% $CaCO_3$ and cells at a final concentration of 10 absorbance units at 600 nanometers (OD_{600}). At the conclusion of the incubation period, the reaction mixtures were analyzed by high performance liquid chromatography (HPLC) using an Agilent 1100 HPLC system (Agilent Technologies, Wilmington, USA) with a LiChrospher-100-RP18 (125 x 4.6 mm) column (Merck, Darmstadt, Germany) attached to an Aminex-HPX-78H (300 x 7.8 mm) column (Biorad, Reinach, Switzerland). The mobile phase was 0.004 M sulfuric acid, and the flow rate was 0.6 ml/min. Two signals were recorded using a UV detector (wavelength 254 nm) in combination with a refractive index detector. In addition, the identification of the L-ascorbic acid was done using an amino-column (YMC-Pack Polyamine-II, YMC, Inc., Kyoto, Japan) with UV detection at 254 nm. The mobile phase was 50 mM $NH_4H_2PO_4$ and acetonitrile (40:60).

An Agilent Series 1100 HPLC-mass spectrometry (MS) system was used to identify L-ascorbic acid. The MS was operated in positive ion mode using the electrospray interface. The separation was carried out using a LUNA-C8(2) column (100 x 4.6 mm) (Phenomenex, Torrance, USA). The mobile phase was a mixture of 0.1% formic acid and methanol (96:4). L-Ascorbic acid eluted with a retention time of 3.1 minutes. The identity of the L-ascorbic acid was confirmed by retention time and the molecular mass of the compound.

The *Gluconobacter* and *Acetobacter* strains produced L-ascorbic acid from L-sorbosone as shown in Table 1.

Table 1. Production of L-ascorbic acid from L-sorbosone

Strain	L-Ascorbic acid produced (mg/L
G. melanogenus IFO 3293	1740
G. melanogenus N44-1	570
G. melanogenus IFO 3292	410
G. rubiginosus IFO 3244	1280
G. industrius IFO 3260	50
G. cerinus IFO 3266	140
G. oxydans IFO 3287	60
A. aceti subsp. orleanus IFO 3259	30
A. aceti subsp. xylinum IFO 13693	40
A. aceti subsp. xylinum IFO 13773	120
A. sp ATCC 15164	310
Blank	not detected

Blank; reaction was done in the reaction mixture without cells

Example 2. L-Ascorbic acid production under various conditions

a) Condition 1. L-Ascorbic acid production from L-sorbosone, L-sorbose or D-sorbitol in a resting cell reaction.

Cells of *G. melanogenus* N44-1 were grown at 27°C for 3 days on No. 3BD agar medium containing 70 g/l L-sorbose, 0.5 g/l glycerol, 7.5 g/l yeast extract (Difco), 2.5 g/l $MgSO_4 \cdot 7H_2O$, 10 g/l $CaCO_3$ and 18 g/l agar (Difco). The resting cell reactions (1 ml reaction mixture in 10 ml tube) were carried out with 2% D-sorbitol, 2% L-sorbose, or 1% L-sorbosone at 30°C for 24 h as described in Example 1. Strain N44-1 produced 280, 400 and 1780 mg/l of L-ascorbic acid from D-sorbitol, L-sorbose, and L-sorbosone, respectively.

Other reactions (0.5 ml reaction mixture in 10 ml tube) were carried out with N44-1 cells grown on No. 3BD agar medium in reaction mixtures containing 2% D-sorbitol, 2% L-sorbose or 2% L-sorbosone for 2 days as described in Example 1. Strain N44-1 produced 1.8, 2.0 and 5.1 g/l of L-ascorbic acid from D-sorbitol, L-sorbose, and L-sorbosone, respectively.

A reaction using cells of *G. melanogenus* IFO 3293 was carried out with 2% L-sorbosone as described above. Strain IFO 3293 produced 5.7 g/l of of L-ascorbic acid in 2 days.

b) Condition 2. L-Ascorbic acid production from L-sorbose and D-sorbitol in tube and flask fermentations.

Cells of *G. melanogenus* N44-1 were used to inoculate 4 ml of No. 3BD liquid medium and cultivated in a tube (18 mm diameter) at 30°C for 3 days with shaking at 220 rpm. Twenty mg/l of L-ascorbic acid had accumulated at the end of the incubation period.

Cells of *G. melanogenus* N44-1 were cultivated (in triplicate) in 50 ml of No. 5 medium containing 100 g/l D-sorbitol, 0.5 g/l glycerol, 15 g/l yeast extract (Difco), 2.5 g/l of $MgSO_4 \cdot 7H_2O$, and 15 g/L of $CaCO_3$ in a 500 ml baffled shake flask at 30°C with shaking at 200 rpm. After 72 h of cultivation, the measured amounts of L-ascorbic acid in the three flasks were 400, 500 and 680 mg/l.

c) Condition 3. Production of L-ascorbic acid from D-sorbitol in batch cultivations.

Cells of *G. melanogenus* IFO 3293 and *G. melanogenus* N44-1 were cultivated in 50 ml No. 5 medium containing 100 g/l D-sorbitol, 0.5 g/l glycerol, 15 g/l yeast extract (Flul BioChemika, Buchs, Switzerland), 2.5 g/l $MgSO_4 \cdot 7H_2O$ and 15 g/l $CaCO_3$ in 500 ml baffle shake flasks at 30°C with shaking at 180 rpm. After 48 h, strains IFO 3293 and N44-1 ha produced 90 mg/l and 60 mg/l L-ascorbic acid, respectively. At 72 h, the concentration of l ascorbic acid produced by N44-1 had increased to 180 mg/l.

d) Condition 4. Production of L-ascorbic acid from D-sorbitol in fed-batch cultivations.

Cells of *G. melanogenus* N44-1 were grown in 200 ml No. 5 medium containing 100 g/l D sorbitol, 0.5 g/l glycerol, 15 g/l yeast extract (Fluka BioChemika, Buchs, Switzerland), 2.5 g/ $MgSO_4 \cdot 7H_2O$ and 15 g/l $CaCO_3$ in a 2l baffled shake flask at 30°C with shaking at 180 rpm After 48 h, 150 ml of this culture was used to inoculate a 10-l bioreactor (B. Braun ED10 Melsungen, Germany) previously prepared with 5.3 l medium containing 20 g/l D-sorbitol, 0.5 g/l glycerol, 15 g/l yeast extract (Fluka BioChemika, Buchs, Switzerland) and 2.5 g/l $MgSO_4.7H_2O$ and equipped with temperature, pH and dissolved oxygen sensors and control loops. Temperature was controlled at 30°C, pH was controlled at 6.0 by adding a 28% ammonia solution, airflow was 4.5 l/min and dissolved oxygen was controlled at 30% by a cascade with stirring speed (minimum 300 rpm). After 6 h process time, a 500 g/l sorbitol solution was fed at a rate of 25 g/h for a period of 44 h. After 96 h process time, about 1% substrate was left in the supernatant, and 950 mg/l L-ascorbic acid had been produced.

e) Condition 5. Production of L-ascorbic acid from D-sorbitol in resting cell mode.

Cells of *G. melanogenus* N44-1 were grown in 200 ml of No. 5 medium containing 100 g/l D-sorbitol, 0.5 g/l glycerol, 15 g/l yeast extract (Fluka BioChemika, Buchs, Switzerland), 2.5 g/l $MgSO_4.7H_2O$ and 15 g/l $CaCO_3$ in a 2 l baffled shake flask at 30°C with shaking at 180 rpm. After 24 h, the culture was centrifuged at 3220 g (Eppendorf 5810R, Hamburg, Germany), and the cells were resuspended in 0.9% NaCl solution, centrifuged again at 3220 g and the cell pellet was used to inoculate one baffled 500 ml shake flask containing 50 ml of full medium (100 g/l D-sorbitol, 0.5 g/l glycerol, 15 g/l yeast extract, 2.5 g/l $MgSO_4.7H_2O$, 15 g/l $CaCO_3$) and another baffled 500 ml shake flask containing 50 ml resting cell medium

(100 g/l D-sorbitol, 3 g/l NaCl, 10 g/l $CaCO_3$). The initial cell density, measured as optical density at 600 nm (OD_{600}), in both flasks was 10. Both flasks were incubated at 30°C with shaking at 180 rpm. After 48 h, the cell suspension in full medium and resting cell medium had accumulated 1.06 and 1.18 g/l L-ascorbic acid, respectively. No additional growth was observed in full medium during the incubation period time.

f) Condition 6. L-Ascorbic acid production from L-sorbosone or L-sorbose with a cell membrane fraction.

Cells of *G. melanogenus* N44-1 were cultivated in 100 ml of No. 3BD liquid medium in a 500 ml baffled shake flask at 30°C with shaking at 220 rpm for 3 days. The resulting culture was centrifuged at 500 rpm to remove $CaCO_3$. The supernatant from this step was then centrifuged at 5,000 rpm to pellet the cells. The collected cells were suspended in 3 ml of 50 mM potassium phosphate buffer (pH7.0) and the cells were disrupted by two passages through a French Pressure cell (SIM-AMINCO Spetronic Instruments, USA) at 900 psi. The resulting homogenate was first centrifuged at 5,000 rpm to remove cell debris. Then the supernatant was diluted to a final protein concentration of 3 mg of protein/ml. This diluted sample is designated as cell-free extract (CFE). The CFE was centrifuged at 100,000 x g for 60 min. The supernatant was discarded and the pellet was collected as the membrane fraction.

The reaction (200 µl) with the membrane fraction (100 µl) was carried out in 50 mM potassium phosphate buffer (pH7.0), 30°C with shaking at 220 rpm for 15 h. The substrates tested and their final concentrations were L-sorbosone (1%) and L-sorbose (2%). The final protein concentration used in the reaction was 1.5 mg/ml. At the end of the incubation period, 680 mg/L and 10 mg/l of L-ascorbic acid had been produced from 1% L-sorbosone and 2% L-sorbose, respectively.

Example 3. Isolation of the SNDHai gene

1. Tn5 mutagenesis

Tn5, a transposable element coding for kanamycin resistance (Km^r), causes null mutations a random locations on the DNA of the host organism and widely is used as a mutagen in Gram negative bacteria. Plasmid pSUP2021, a "suicide" vector containing Tn5 (Simon R. et al. 198 BIO/TECHNOLOGY 1: 784-791), was transferred from *E. coli* HB101 into *G. melanogenu* N44-1 by a conjugal mating method as follows. *G. melanogenus* N44-1 was cultivated in a tes tube containing 5 ml of MB liquid medium at 30°C overnight. *E. coli* HB101 carrying helpe plasmid pRK2013 [D. H. Figurski, Proc. Natl. Acad. Sci. USA, 76, 1648-1652, 1979]) and *E. co* HB101 carrying plasmid pSUP2021 were cultivated in test tubes containing 5 ml of LB mediun with 50 μg/ml of kanamycin at 37°C overnight. From the overnight cultures, cells of *G melanogenus* N44-1, *E. coli* HB101(pRK2013), and *E. coli* HB101(pSUP2021) were collectec separately by centrifugation and suspended to the original volume in MB medium. Then these cell suspensions were mixed in equal volumes and the mixture was spread out on a 0.45 μm nitrocellulose membrane laid on top of an MB agar plate. After cultivation at 27°C for one day, the cells were scraped off the membrane and dilutions were prepared in MB broth. The diluted cells were then spread on MB agar medium containing 10 μg/ml of polymixin B and 50 μg/ml of kanamycin (MPK medium). The polymixin B selects against the *E. coli* donor and helper strains, while the kanaymycin selects for those *G. melanogenus* cells that have been transformed with plasmid pSUP2021 (i.e., the transconjugants). About 30,000 transconjugants were obtained.

2. Screening for L-ascorbic acid non-producers.

In all, 3,760 transconjugants were transferred with sterile toothpicks onto MPK grid plates and grown at 27°C for 3 days. To test for L-ascorbic acid production from L-sorbosone, cells of each transconjugant were picked off the grid plate with a sterile toothpick and suspended in 50 μl of a resting cell reaction mixture containing 0.5% L-sorbosone, 0.3% NaCl, and 1% $CaCO_3$ in 96-well microtiter plates. The microtiter plates were incubated at 30°C for one day without shaking. One microliter of each of the resulting reaction mixtures was analyzed for L-ascorbic acid formation using ascorbic acid test strips and the RQFlex2 instrument (Merck KGaA, 64271 Darmstadt, Germany). The positive control strain was *G. melanogenus* N44-1

grown under identical conditions. By this method, the L-ascorbic acid-non-producing mutant N44-1-6A9 was identified. Southern blot hybridization analysis was then performed to confirm the presence of Tn5 in the chromosomal DNA of mutant N44-1-6A9. Two µg of chromosomal DNA isolated from the mutant was digested with either *Apa*I, *Cla*I, *Eco*RI, *Eco*RV, *Kpn*I, *Stu*I, *Bam*HI, *Sal*I, or *Hin*dIII and subjected to agarose gel electrophoresis (0.8% agarose). The gel was then treated with 0.25 N HCl for 15 min, followed by 0.5 N NaOH for 30 min. The DNA was transferred to a nylon membrane with the rapid downward transfer system TurboBlotter (Schleicher & Schuell GmbH, Germany). The probe was prepared with PCR-DIG labeling kit (Roche Diagnostics GmbH, 68298 Mannheim, Germany) using primers Tn2419 (5'- CGCCTTCTATGAAAGGTTGG-3'; SEQ ID NO:3) and Tn3125R (5'-AGCGGATGGAGATCGGGCGG-3'; SEQ ID NO:4) with plasmid pSUP2021 as the template. A 707-bp PCR product was obtained.

The hybridization conditions used were as follows:

Hybridization was done under stringent hybridization conditions, e.g., 2 hours to 4 days incubation at 42°C using a digoxigenin (DIG)-labeled DNA probe (prepared by using a DIG labeling system; Roche Diagnostics GmbH, 68298 Mannheim, Germany) in DigEasyHyb solution (Roche Diagnostics) with 100 µg/ml salmon sperm DNA, followed by washing the filters for 15 min (twice) in 2 x SSC and 0.1% SDS at room temperature and then washing for 15 min (twice) in 0.5 x SSC and 0.1% SDS or 0.1 x SSC and 0.1% SDS at 65°C.

Following the hybridization step, the detection of hybridization was done with anti-DIG-AP conjugate (Roche Diagnostics GmbH, 68298 Mannheim, Germany) and ECF substrate (Amersham Biosciences Uppsala, Sweden) using the STORM instrument (Amersham Biosciences). All operations were done according to the instructions of the suppliers.

Using the methods described above, the presence of Tn5 in the chromosome of mutant N44-1-6A9 was confirmed.

3. Cloning of a DNA fragment interrupted by Tn5 and sequencing of the adjacent regions.

Based on the results of the restriction enzyme digestions described in section 2 above, *Apa* *Cla*I, *Eco*RI, and *Eco*RV were selected as the enzymes that generated DNA fragment havin more than 1 kb of flanking chromosomal DNA at both sides of the Tn5 insertion. Doubl digestion of mutant N44-1-6A9 DNA with *Sal*I (which cuts approximately in the middle of Tn5 and *Apa*I gave two fragments (6.2 and 3.8 kb) that hybridized to the 707-bp probe describe above.

The chromosomal DNA of the Tn5 mutant *G. melanogenus* N44-1-6A9 was prepared anc digested with *Apa*I. The DNA fragments with the size of from 9 to 12 kb were isolated from agarose gel and ligated with the cloning vector pBluescript II KS+ (Stratagene, Switzerland), previously digested with *Apa*I. The ligation mixture was then used to transform competent *E. coli* cells, selecting on L-agar plates containing 50µg/ml kanamycin and 100µg/ml ampicillin. From one transformant, the plasmid was extracted and the cloned regions flanking the Tn5 insertion were sequenced. The nucleotide sequence of the single open reading frame (interrupted by the Tn5 insertion) was assembled after removing a 9-bp duplication that is known to occur during the Tn5 transposition event. The nucleotide sequence of the full open reading frame, hereafter referred to as the SNDHai gene of *Gluconobacter melanogenus* N44-1, consisted of 2367 bp and is given as SEQ ID NO:1. The corresponding amino acid sequence deduced from the nucleotide sequence of SEQ ID NO:1 is given here as SEQ ID NO:2.

The nucleotide sequence of the SNDHai gene (SEQ ID NO:1) was subjected to a Blast 2 search (version 2 of BLAST from the National Center for Biotechnology Information [NCBI] described in Altschul et al. (Nucleic Acids Res. 25: 3389-3402, 1997)) on the database PRO SW-SwissProt (full release plus incremental updates). The conditions used were the gapped alignment and filtration of the query sequence for the low complexity region. The corresponding amino acid sequence of the full length SNDHai gene (SEQ ID NO:2) showed moderate homologies (36-40%) with membrane-bound PQQ-dependent D-glucose dehydrogenases and membrane-bound PQQ-dependent quinate/shikimate dehydrogenases (Table 2).

Table 2. Homology of SNDHai and other dehydrogenases

Enzyme Name	Accession No.	Total (residues)	Identity
SNDHai	-	788	100%
Quinate/shikimate DH	Q59086	809	40%
Probable quinate DH	Q9xd78	790	39%
Glucose DH	P15877	796	36%
Glucose DH	P05465	801	36%
Glucose DH	P27175	808	36%

Analysis of the structure of the SNDHai gene revealed an N-domain and a C-domain that indicate a structure of the membrane-spanning region (140-150 residues, hydrophobic) and a primary dehydrogenase region (650-660 residues), respectively, designated according to the membrane-bound D-glucose dehydrogenase described by Yamada et al. (J Biol. Chem., 276:48356-48361, 2001). Besides the sequences listed in Table 2, the SNDHai showed some lower homologies (26-28%) within the C-domain with the other quinoproteins such as alcohol dehydrogenases of *Acetobacter* and *Gluconobacter* (accession Nos. Q44002 and O05542, respectively), and methanol dehydrogenases of *Methylobacterium* (P16027 and P15279), *Methylotrophus* (P38539) and *Paracoccus* (P12293).

When the amino acid sequence of SEQ ID NO:2 was analyzed using the program MOTIFS, bacterial quinoprotein dehydrogenases signature sequences were readily identified.

Bacterial_Pqq_2 sequence (Prosite code: PS00364) was identified as follows:

Bacterial_Pqq_2 motif:

Wx4YDx3(D,N)(L,I,V,M,F,Y)4x2Gx2(S,T,A)PWx{4}YDx{3}(N)(L,V,F){4}x{2}Gx{2}(T)P

SNDHai from amino acid 396:

WGTASYDPKLNLVFFPLGNQTP

Bacterial_Pqq_1 sequence (Prosite code: PS00364) was also identified, but with one differenc in the amino acids compared to SNDHai (underligned A in sequence below):

Bacterial_Pqq_1 Motif:

(DN)Wx3G(RK)x6(FY)Sx4(LIVM)Nx2NVx2L(RK)

SNDHai from amino acid 168:

DWPAYGRTASGTRYASFTQINRDNVSKLR

Identification of these motifs indicates that SNDHai has the characteristics of a PQQ-dependent enzyme.

When the amino acid sequence of SNDHai was compared to protein sequences in the database PRO GSP-Gene Seq patented protein sequences, the peptides with higher E value (e^{-110} to $6e^{-12}$) were sorbitol dehydrogenases (accession numbers Aaw95019 and Aab35987), alcohol dehydrogenases (Aar20192 and Aar13993), and alcohol and/or aldehyde dehydrogenases (Aaw37876, Aaw37875, Aaw37873, and Aaw37874), while those peptides with lower E value ($9e^{-13}$ to 1.7) were cyclitol-ubiquinone-oxidoreductase peptides (Aaw36473, Aaw36460, Aaw36465, Aaw36462, Aaw36466, Aaw36459, Aaw36471, Aaw36455, Aaw36474, Aaw36454, Aaw36470 and Aaw36468).

Example 4. Southern blot analysis of the bacteria producing L-ascorbic acid from L-sorbosone

Chromosomal DNA was prepared from cells of *Gluconobacter melanogenus* IFO 3293, *Gluconobacter melanogenus* IFO 3292, *Gluconobacter rubiginosus* IFO 3244, *Gluconobacter industrius* IFO 3260, *Gluconobacter cerinus* IFO 3265, *Gluconobacter cerinus* IFO 3266, *Gluconobacter cerinus* IFO 3269, *Gluconobacter oxydans* IFO 3287, *Acetobacter aceti subsp. orleanus* IFO 3259, *Acetobacter aceti subsp. xylinum* IFO 13693, *Acetobacter aceti subsp. xylinum* IFO 13773, *Acetobacter sp.* ATCC 15164, and *Escherichia coli* K-12. All *Acetobacter* and *Gluconobacter* strains were grown on MB agar medium as described in Example 1. *E. coli* K-12 was grown on Luria Broth agar medium. The chromosomal DNA preparations were used for Southern blot hybridization under stringent conditions as described in Example 3. The chromosomal DNA preparations were digested with *Cla*I (when analyzing the N-domain region) or *Eco*RI (when analyzing the C-domain region), and 1 μg of the DNA fragments were separated by agarose gel electrophoresis (1% agarose). The gel was treated with 0.25 N HCl for 15 min and then 0.5 N NaOH for 30 min, and then was blotted onto a nylon membrane with Vacuum Blotter Model 785 (BIO-RAD Laboratories AG, Switzerland) according to the instruction of the supplier. The probes were prepared with PCR-DIG labeling kit (Roche Diagnostics) by using the primer sets for PCR products P1 and P2 (described in Table 3). The PCR product P1 corresponds to the region of SNDHai designated the N-domain (possible transmembrane region) while PCR product P2 corresponds to the region of SNDHai designated as the C-domain (possible primary dehydrogenase region).

Table 3. Primers used for PCR to generate labeled probes for Southern hybridizations.

Primer set and primer names	Sequence	Expected size of PCR product
Primer set for PCR product P1		420 bp
SNDH1F (SEQIDNO:5)	5'-ATGAACAGCGGCCCCCGCACGCTCTCCATG-3'	
SNDH420R (SEQIDNO:6)	5'-CCGGAACATGCCGGCGAGGAAAGCCACGAT-3'	
Primer set for PCR product P2		1864 bp
SNDH501F (SEQIDNO:7)	5'-TGACTGGCCCGCCTATGGCCGCACGGCTTC-3'	
SNDH2364R (SEQIDNO:8)	5'-TTCTTCGGAGGGCAGGGCGTAGGCGATGAC-3'	
Primer set for PCR product P3		1030 bp
SNDH501F(SEQIDNO:7)	5'-TGACTGGCCCGCCTATGGCCGCACGGCTTC-3'	
SNDH1530R(SEQIDNO:9)	5'-CGGGACTTTGCGCATTTCCACAGGGACGAT-3'	
Primer set for PCR product P4		974 bp
1391F(SEQIDNO:10)	5'-AGCCCATCCTCTATGACATTCCGGACGGCC-3'	
2364R(SEQIDNO:8)	5'-TTCTTCGGAGGGCAGGGCGTAGGCGATGAC-3'	

Table 4 shows the results of the Southern blot hybridization experiments. In the hybridizatic with the P1 (N-domain) probe (Table 4), clear positive bands were observed f *G. melanogenus* IFO 3293, *G. melanogenus* IFO 3292, *G rubiginosus* IFO 3244 *G. oxydar* IFO 3287 and *A.* sp. ATCC 15164. In the hybridization with the P2 (C-domain) probe, cle

positive bands were observed for *G. melanogenus* IFO 3293, *G. melanogenus* IFO 3292, and *G. rubiginosus* IFO 3244, *G. oxydans* IFO 3287 and *A. sp.* ATCC 15164, while a faint band was observed for *G. industrius* IFO 3260, *G. cerinus* IFO 3265, *G. cerinus* IFO 3266, *G. cerinus* IFO 3269 and *A. aceti subsp. xylinum* IFO 13773. The control strain, *E. coli* K-12, showed no detectable signals for either domain.

Table 4 Southern blot hybridization with probes for N- and C-domains of SNDHa

Strain	Hybridization signal with probe generated by PCR primer se	
	P1 (corresponds to N-domain)	P2 (corresponds to C-domair
G. melanogenus IFO 3293	+	+
G. melanogenus IFO 3292	+	+
G. rubiginosus IFO 3244	+	+
G. industrius IFO 3260	nd	tr
G. cerinus IFO 3265	nd	tr
G. cerinus IFO 3266	nd	tr
G. cerinus IFO 3269	nd	tr
G. oxydans IFO 3287	+	+
A. aceti subsp. orleanus IFO 3259	nd	nd
A. aceti subsp. xylinum IFO 13693	nd	nd
A. aceti subsp. xylinum IFO 13773	nd	tr
A. sp. ATCC 15164	+	+
E. coli K-12	nd	nd

Tr, trace; nd, not detected. Probes P1 and P2 were synthesized (as DIG-labeled PCR products) with the primer sets specified in Table 3.

Example 5. PCR amplification and sequencing of orthologs of the SNDHai gene

Chromosomal DNA preparations (prepared as described in Example 4) were used as templates for PCR with the four primer sets shown in Table 3. Five to 100 ng of chromosomal DNA was used per reaction (total volume, 50 µl). Unless specified otherwise, the Expand High Fidelity PCR system was used (Roche Diagnostics). The PCR conditions were as follows:

1. Incubation at 94°C for 2 min.

2. Thirty cycles of (i) denaturation step at 94°C for 15 sec; (ii) annealing step at 60°C for 30 sec; (iii) synthesis step at 72°C for 45 to 120 sec (time for the synthesis step for primer sets P1, P2, P3 and P4 were 45 sec, 120 sec, 90 sec, and 90 sec, respectively).

3. Extension at 72°C for 7 min.

Samples of the PCR reactions were separated by agarose gel electrophoresis and the bands were visualized with a transilluminator after staining with ethidium bromide. The results of the PCR reactions are summarized in Table 5.

Table 5

Strain	PCR product			
	Primer set for PCR product	P1	P2	P3	P4
G. melanogenus IFO 3293		+	+*	nt	+
G. melanogenus IFO 3292		+	nd	nd	+
G. rubiginosus IFO 3244		+	+	+	+
G. industrius IFO 3260		nd	nd	nd	nd
G. cerinus IFO 3266		nd	nd	nd	nd
G. oxydans IFO 3287		+	+	nd	+
A. aceti subsp. orleanus IFO 3259		nd	nd	nd	nd
A. aceti subsp. xylinum IFO 13693.		nd	nd	nd	nd
A. aceti subsp. xylinum IFO 13773		nd	nd	nd	nd
A. sp ATCC 15164		+	+	nd	nd
E. coli K-12		nd	nd	nt	nd

+, detected; nd, not detected; nt, not tested.

*This PCR was done with GC-rich PCR system (Roche Diagnostics) with the same reaction cycle as was used for Expand High Fidelity PCR system.

Comparison of the nucleotide and amino acid sequence of SNDHai from *G. melanogenus* N44-1 with PCR products.

When clear PCR bands were observed on the agarose gel (Table 5), the PCR products were used directly for nucleotide sequencing using standard methods. The nucleotide sequences obtained for the different PCR products, and the corresponding amino acid sequences of the encoded peptides, were compared with the full length sequence of the SNDHai gene and protein from *G. melanogenus* N44-1. An alignment of selected amino acid sequences is shown in Fig. 1.

Gluconobacter melanogenus IFO 3292

The PCR product (about 1 kb) obtained upon amplification with primers SNDH1391F (SEQIDNO:10) and SNDH2364R (SEQIDNO:8) and chromosomal DNA from *G. melanogenus* IFO 3292 as the template, was used for sequencing with primer SNDH1391F (SEQIDNO:10). The determined nucleotide sequence (SEQ ID NO:11, 771 bp) showed 98.7% (761/771) identity with nucleotides 1431-2201 of the sequence of SNDHai from *G. melanogenus* N44-1 (nucleotides 1431-2201 of SEQ ID NO:1). The deduced amino acid sequence (SEQ ID NO:12, 256 amino acids) showed 100% identity to amino acids 478-733 of the amino acid sequence of SNDH from *G. melanogenus* N44-1 (amino acids 478-733 of SEQ ID NO:2).

Gluconobacter rubiginosus IFO 3244

The PCR product (about 0.4 kb) obtained upon amplification with primers SNDH1F (SEQ ID NO:5) and SNDH420R (SEQ ID NO:6) and chromosomal DNA from *G. rubiginosus* IFO 3244 as the template, was used for sequencing with primer SNDH420R (SEQ ID NO:6). The determined nucleotide sequence (SEQ ID NO:13, 344 bp) showed 98.5% (339/344) identity with the nucleotides 31-374 of the sequence of SNDHai from *G. melanogenus* N44-1 (nucleotides 31-374 of SEQ ID NO:1). The deduced amino acid sequence (SEQ ID NO:14, 114 residues) showed 100% identity to amino acids 11-124 of the amino acid sequence of SNDHai from *G. melanogenus* N44-1 (amino acids 11-124 of SEQ ID NO:2).

The PCR product (about 1 kb) obtained upon amplification with primers SNDH501F (SEQ NO:7) and 1530R (SEQ ID NO:9) was used for sequencing with primer SNDH153((SEQIDNO:9). The determined nucleotide sequence (SEQ ID NO:15, 110 bp) showed 100 (110/110) identity with the nucleotides 1391-1500 of the nucleotide sequence of SNDHai fro *G. melanogenus* N44-1 (nucleotides 1391-1500 of SEQ ID NO:1). The deduced amino ac sequence (SEQ ID NO:16, 36 residues) showed 100% identity to amino acids 465-500 of th amino acid sequence of SNDHai from *G. melanogenus* N44-1 (amino acids 465 to 500 of SE ID NO:2).

The PCR product (about 1 kb) obtained upon amplification with primers SNDH1391F (SEQ II NO:10) and SNDH2364R (SEQ ID NO:8) was used for sequencing with primer SNDH1391I (SEQ ID NO:10). The determined nucleotide sequence (SEQ ID NO:17, 801 bp) showe 98.9% (792/801) identity with nucleotides 1458-2258 of the nucleotide sequence of SNDHa from *G. melanogenus* N44-1 (nucleotides 1458-2258 of SEQ ID NO:1). The deduced amin acid sequence (SEQ ID NO:18, 266 residues) showed 100% identity with amino acids 487-75 of the amino acid sequence of SNDHai from *G. melanogenus* N44-1 (amino acids 487-752 o SEQ ID NO:2).

Gluconobacter oxydans IFO 3287

The PCR product (about 0.4 kb) obtained upon amplification with primers SNDH1F (SEQ ID NO:5) and SNDH420R (SEQ ID NO:6) and chromosomal DNA from *G. oxydans* IFO 3287 as the template, was used for sequencing with primer SNDH420R (SEQ ID NO:6). The determined nucleotide sequence (SEQ ID NO:19, 350 bp) showed 97.4% (341/350) identity with nucleotides 31-380 of the nucleotide sequence of SNDHai from *G. melanogenus* N44-1 (nucleotides 31-380 of SEQ ID NO:1). The deduced amino acid sequence (SEQ ID NO:20, 116 amino acids) showed 100% identity with amino acids 11-126 of the amino acid sequence of SNDHai from from *G. melanogenus* N44-1 (amino acids 11-126 of SEQ ID NO:2).

The PCR product (about 1.9 kb) obtained upon amplification with primers SNDH501F (SEQ ID NO:7) and SNDH2364R (SEQ ID NO:8) was used for sequencing with primer SNDH501F (SEQ ID NO:7). The determined nucleotide sequence (SEQ ID NO:21, 808 bp) showed 98.0%

(745/808) identity with nucleotides 578-1385 of the nucleotide sequence of SNDHai from *G. melanogenus* N44-1 (nucleotides 578-1385 of SEQ ID NO:1). The deduced amino acid sequence (SEQ ID NO:22, 268 residues) showed 100% identity to amino acids 194-461 of the amino acid sequence of SNDHai from *G. melanogenus* N44-1 (amino acids 194-461 of SEQ ID NO:2).

The PCR product (about 1 kb) obtained upon amplification with primers SNDH1391F (SEQ ID NO:10) and SNDH2364R (SEQ ID NO:8) was used for sequencing with primer SNDH1391F (SEQ ID NO:10). The determined nucleotide sequence (SEQ ID NO:23, 800 bp) showed 98.8% (790/800) identity with nucleotides 1469-2268 of the nucleotide sequence of SNDHai from *G. melanogenus* N44-1 (nucleotides 1469-2268 of SEQ ID NO:1). The deduced amino acid sequence (SEQ ID NO:24, 266 residues) showed 100% identity with amino acids 491-756 of the amino acid sequence of SNDHai from *G. melanogenus* N44-1 (amino acids 491-756 of SEQ ID NO:2).

Acetobacter sp. ATCC 15164

The PCR product (about 0.4 kb) obtained upon amplification with primers SNDH1F (SEQ IDNO:5) and SNDH420R (SEQ ID NO:6) and chromosomal DNA from *A. sp.* ATCC 15164 as the template, was used for sequencing with primer SNDH420R (SEQ ID NO:6). The determined nucleotide sequence (SEQ ID NO:25, 360 bp) showed 97.8% (352/360) identity with nucleotides 31-390 of the nucleotide sequence of SNDHai from *G. melanogenus* N44-1 (nucleotides 31-390 of SEQ IDNO:1). The deduced amino acid sequence (SEQ ID NO:26, 120 residues) showed 100% identity with amino acids 11-130 of the amino acid sequence of SNDHai from *G. melanogenus* N44-1 (amino acids 11-130 of SEQ ID NO:2).

The PCR product (about 1.9 kb) obtained upon amplification with primers SNDH501F (SEQ ID NO:7) and SNDH2364R (SEQ ID NO:8) was used for sequencing with primer SNDH501F (SEQ ID NO:7). The determined nucleotide sequence (SEQ ID NO:27, 760 bp) showed 98.0% (745/760) identity with nucleotides 563-1322 of the nucleotide sequence of SNDHai from *G. melanogenus* N44-1 (nucleotides 563-1322 of SEQ ID NO:1). The deduced amino acid sequence (SEQ ID NO:28, 252 residues) showed 100% identity with amino acids 189-440 of the amino acid sequence of SNDHau from *G. melanogenus* N44-1 (amino acids 189-440 of SEQ ID NO:2).

Example 6

Increased L-ascorbic acid production from L-sorbosone by increasing the SNDHai ger dosage

The SNDHai gene with upstream and downstream flanking regions was amplified by PCR wit chromosomal DNA of strain N44-1 as template and the following primers:

SEQ ID NO:29 5'-CCGAATTCAGGCCGAACAGCAGCAGGTCAC

SEQ ID NO:30 5'-GTGCCTGGGTACCTCGGTGGAGGTCATGAA.

The PCR was done with the GC-rich PCR system (Roche Diagnostics) according to the instructions of the supplier. The amplified DNA fragment was inserted into vector pCR2.1-TOPO. The resulting plasmid was then digested with *Hin*dIII and *Xho*I. The *Hin*dIII-*Xho*I fragment including the SNDHai gene was ligated to vector pVK100 (available from the American Type Culture Collection, catalog no. ATCC 37156) previously treated with *Hin*dIII and *Xho*I. The ligation mixture was used to transform *E.coli* TG1. The desired plasmid, designated pVK-P-SNDHai-T, was isolated from *E.coli*, and introduced into *G.melanogenus* strain N44-1 by electroporation using standard methods (Electrocell manipulator ECM600, BTX Inc., San Diego, CA, USA).

Three independent transformants, designated N44-1 (pVK-P-SNDHai-T) clone number 1, 2, and 3, together with the parental strain *G.melanogenus* N44-1, were each grown on No. 3BD agar and MB agar media. The cells were scraped from the plates and used for resting cell reactions (1% L-sorbosone as the substrate) as described in Example 1. After growth on No. 3BD agar, in the resting cell assay strain N44-1 produced 2.5 g/l L-ascorbic acid, while strains N44-1 (pVK-P-SNDHai-T) clones 1, 2 and 3 produced 4.2, 4.1 and 4.2 g/l L-ascorbic acid, respectively. After growth on MB agar, in the resting cell assay strain N44-1 produced 0.12 g/l L-ascorbic acid, while strains N44-1 (pVK-P-SNDHai-T) clones 1, 2 and 3 produced 1.8, 2.5 and 0.94 g/l L-ascorbic acid, respectively.

Case 21864

Claims

1. An isolated polynucleotide molecule derived from a polynucleotide encoding a polypeptic having L-sorbosone dehydrogenase activity comprising a partial nucleotide sequence of : least 20 consecutive nucleotides of SEQ ID No:1.

2. Isolated polynucleotide molecule according to claim 1, wherein the partial nucleotid sequence of SEQ ID No:1 has at least 50 consecutive nucleotides.

3. Isolated polynucleotide molecule according to claim 1, wherein the partial nucleotid sequence of SEQ ID No:1 has at least 100 consecutive nucleotides.

4. Isolated polynucleotide according to claim 3 wherein the partial nucleotide sequence i derived from a polynucleotide sequence having a homology of at least 60% identity with SEC ID No:1 whereby at least 100 consecutive nucleotides are compared.

5. Isolated polynucleotide molecule according to any of the preceding claims, whereby the partial nucleotide sequence is derived from a polynucleotide sequence having a homology of at least 80% identitiy with SEQ ID No:1.

6. Isolated polynucleotide molecule according to any of the preceding claims, whereby the partial nucleotide sequence is derived from a polynucleotide sequence having a homology of at least 90% identity with SEQ ID No:1.

7. Isolated polynucleotide molecule according to claim 1, wherein the partial nucleotide sequence is selected from the group consisting of SEQ ID Nos:5, 6, 7, 8, 9 and 10.

8. Isolated polynucleotide molecule according to any of claims 1-6, which is selected from the group consisting of SEQ Nos:1, 11, 13, 15, 17, 19, 21, 23, 25 and 27.

9. Polypeptide characterized in that it is encoded by a polynucleotide according to any of the preceding claims.

10. Polypeptide according to claim 9, comprising a partial amino acid sequence of at least 25 consecutive amino acids selected from the group consisting of SEQ ID No:2, 12, 14, 16, 18, 20, 22, 24, 26 and 28.

11. Polypeptide according to claims 9 or 10, wherein the partial amino acid sequence has at least 35 consecutive amino acids.

12. A recombinant DNA molecule for expression of a polypeptide having L-sorbosone dehydrogenase activity characterized in that comprises a polynucleotide according to any of claims 1-6 and 8.

13. Expression vector comprising a recombinant DNA molecule according to claim 12.

14. Recombinant microorganism which has been transformed with a recombinant DNA according to claim 12 and/or an expression vector of claim 13.

15. Recombinant microorganism according to claim 14, wherein the recombinant DNA is at least partially integrated into the chromosome.

16. Recombinant microorganism according to claim 14 or 15, which is selected from the group consisting of fungal, plant, yeast and bacterial cells.

17. Recombinant microorganism according to claim 16, wherein the microorganism is a bacterium of a genus selected from the group consisting of *Gluconobacter, Acetobacter*, *Pseudomonas* and *Escherichia*.

18. A process for the production of L-ascorbic acid from L-sorbosone comprising

a) propagating a recombinant microoganism of any of claims 14-17 in an appropriate culture medium and

b) recovering and separating L-ascorbic acid from said culture medium.

19. A process for the production of L-sorbosone dehydrogenase, wherein a recombinant organism comprising a polynucleotide according to any of claims 1-8 is propagated in an appropriate culture medium, the cells are disrupted and the L-sorbosone dehydrogenase is isolated.

Case 21864

Abstract

The present invention discloses an isolated polynucleotide molecule derived from a polynucleotide encoding a polypeptide having L-sorbosone dehydrogenase activity comprising a partial nucleotide sequence of at least 20 consecutive nucleotides of SEQ ID No:1.

```
N44-1 (SEQIDNO:2)     1  MNSGPRTLSM IIGILGALMA AFLIIEGLHL IILGGSWFYT LAGIALAASS  50
3244  (SEQIDNO:14)                IIGILGALMA AFLIIEGLHL IILGGSWFYT LAGIALAASS
3287  (SEQIDNO:20)                IIGILGALMA AFLIIEGLHL IILGGSWFYT LAGIALAASS
15164 (SEQIDNO:26)                IIGILGALMA AFLIIEGLHL IILGGSWFYT LAGIALAASS

N44-1 (SEQIDNO:2)    51 VYMIRRNILS TWIALGLLVA TALWSLAEVG TSFWPSFSRL IVFLCVALIA 100
3244  (SEQIDNO:14)       VYMIRRNILS TWIALGLLVA TALWSLAEVG TSFWPSFSRL IVFLCVALIA
3287  (SEQIDNO:20)       VYMIRRNILS TWIALGLLVA TALWSLAEVG TSFWPSFSRL IVFLCVALIA
15164 (SEQIDNO:26)       VYMIRRNILS TWIALGLLVA TALWSLAEVG TSFWPSFSRL IVFLCVALIA

N44-1 (SEQIDNO:2)   101 TLMAPWLSGP GRRYFTRPVT GATSGALGAI IVAFLAGMFR VHPTIAPQDT 150
3244  (SEQIDNO:14)       TLMAPWLSGP GRRYFTRPVT GATS
3287  (SEQIDNO:20)       TLMAPWLSGP GRRYFTRPVT GATSGA
15164 (SEQIDNO:26)       TLMAPWLSGP GRRYFTRPVT GATSGALGAI

N44-1 (SEQIDNO:2)   151 THPQETASTA DSDQPGHDWP AYGRTASGTR YASFTQINRD NVSKLRVAWT
3287  (SEQIDNO:22)                                                 KLRVAWT
15164 (SEQIDNO:28)                                               RD NVSKLRVAWT

N44-1 (SEQIDNO:2)   201 YRTGDMALNG AEFQGTPIKI GDTVYICSPH NIVSALDPDT GTEKWKFDPH
3287  (SEQIDNO:22)       YRTGDMALNG AEFQGTPIKI GDTVYICSPH NIVSALDPDT GTEKWKFDPH
15164 (SEQIDNO:28)       YRTGDMALNG AEFQGTPIKI GDTVYICSPH NIVSALDPDT GTEKWKFDPH

N44-1 (SEQIDNO:2)   251 AQTKVWQRCR GVGYWHDSTA TDANAPCASR IVLTTIDARL ITIDARTGQA
3287  (SEQIDNO:22)       AQTKVWQRCR GVGYWHDSTA TDANAPCASR IVLTTIDARL ITIDARTGQA
15164 (SEQIDNO:28)       AQTKVWQRCR GVGYWHDSTA TDANAPCASR IVLTTIDARL ITIDARTGQA

N44-1 (SEQIDNO:2)   301 CTDFGTNGNV NLLTGLGPTA PGSYYPTAAP LVAGDIVVVG GRIADNERTG
3287  (SEQIDNO:22)       CTDFGTNGNV NLLTGLGPTA PGSYYPTAAP LVAGDIVVV GGRIADNERTG
15164 (SEQIDNO:28)       CTDFGTNGNV NLLTGLGPTA PGSYYPTAAP LVAGDIVVV GGRIADNERTG

N44-1 (SEQIDNO:2)   351 EPSGVVRGYD VRTGAQVWAW DATNPHRGTT PLAEGEIYPA ETPNMWGTAS
3287  (SEQIDNO:22)       EPSGVVRGYD VRTGAQVWAW DATNPHRGTT PLAEGEIYPA ETPNMWGTAS
15164 (SEQIDNO:28)       EPSGVVRGYD VRTGAQVWAW DATNPHRGTT PLAEGEIYPA ETPNMWGTAS


N44-1 (SEQIDNO:2)   401 YDPKLNLVFF PLGNQTPDFW GGDRSKASDE YNDAFVAVDA KTGDERWHFR
3287  (SEQIDNO:22)       YDPKLNLVFF PLGNQTPDFW GGDRSKASDE YNDAFVAVDA KTGDERWHFR
15164 (SEQIDNO:28)       YDPKLNLVFF PLGNQTPDFW GGDRSKASDE YNDAFVAVDA

N44-1(SEQIDNO:2)    451 TANHDLVDYD ATAQPILYDI PDGHGGTRPA IIAMTKRGQI FVLDRRDGTP
```

Fig. 1-1

```
3244 (SEQIDNO:16/18)                    PILYDI PDGHGGTRPA IIAMTKRGQI FVLDRRDGTP
3287 (SEQIDNO:22/24)    TANHDLVDYD A                                 FVLDRRDGTP
3292 (SEQIDNO:12)                                        RPA IIAMTKRGQI FVLDRRDGTP

N44-1(SEQIDNO:2)   501  IVPVEMRKVP QDGAPEHQYL APEQPYSALS IGTERLKPSD MWGGTIFDQL
3244 (SEQIDNO:18)       IVPVEMRKVP QDGAPEHQYL APEQPYSALS IGTERLKPSD MWGGTIFDQL
3287 (SEQIDNO:24)       IVPVEMRKVP QDGAPEHQYL APEQPYSALS IGTERLKPSD MWGGTIFDQL
3292 (SEQIDNO:12)       IVPVEMRKVP QDGAPEHQYL APEQPYSALS IGTERLKPSD MWGGTIFDQL

N44-1(SEQIDNO:2)   551  LCRIQFASYR YEGEFTPVNE KQATIIYPGY YGGINWGGGA VDESTGTLLV
3244 (SEQIDNO:18)       LCRIQFASYR YEGEFTPVNE KQATIIYPGY YGGINWGGGA VDESTGTLLV
3287 (SEQIDNO:24)       LCRIQFASYR YEGEFTPVNE KQATIIYPGY YGGINWGGGA VDESTGTLLV
3292 (SEQIDNO:12)       LCRIQFASYR YEGEFTPVNE KQATIIYPGY YGGINWGGGA VDESTGTLLV

N44-1(SEQIDNO:2)   601  NDIRMAQWGK FMKQEEARRS GFKPSSEGEY SEQKGTPWGV VRSMFFSPAG
3244 (SEQIDNO:18)       NDIRMAQWGK FMKQEEARRS GFKPSSEGEY SEQKGTPWGV VRSMFFSPAG
3287 (SEQIDNO:24)       NDIRMAQWGK FMKQEEARRS GFKPSSEGEY SEQKGTPWGV VRSMFFSPAG
3292 (SEQIDNO:12)       NDIRMAQWGK FMKQEEARRS GFKPSSEGEY SEQKGTPWGV VRSMFFSPAG

N44-1(SEQIDNO:2)   651  LPCVKPPYGT MNAIDLRSGK VKWSMPLGTI QDMPVHGMVP GLAIPLGMPT
3244 (SEQIDNO:18)       LPCVKPPYGT MNAIDLRSGK VKWSMPLGTI QDMPVHGMVP GLAIPLGMPT
3287 (SEQIDNO:24)       LPCVKPPYGT MNAIDLRSGK VKWSMPLGTI QDMPVHGMVP GLAIPLGMPT
3292 (SEQIDNO:12)       LPCVKPPYGT MNAIDLRSGK VKWSMPLGTI QDMPVHGMVP GLAIPLGMPT

N44-1(SEQIDNO:2)   701  MSGPLATHTG LVFFSGTLDN YVRALNTDTG EVVWKARLPV ASQAAPMSYM
3244 (SEQIDNO:18)       MSGPLATHTG LVFFSGTLDN YVRALNTDTG EVVWKARLPV ASQAAPMSYM
3287 (SEQIDNO:24)       MSGPLATHTG LVFFSGTLDN YVRALNTDTG EVVWKARLPV ASQAAPMSYM
3292 (SEQIDNO:12)       MSGPLATHTG LVFFSGTLDN YVRALNTDTG EVV

N44-1(SEQIDNO:2)   751  SDKTGKQYIV VTAGGLTRSG VDKNRGDYVI AYALPSEE
3244 (SEQIDNO:18)       SD
3287 (SEQIDNO:24)       SDKTGK
```

Fig. 1-2

SEQUENCE LISTING

<110> Roche Vitamins AG

<120> Microbial Production of L-Ascorbic Acid

<130> 21864

<140>
<141>

<160> 30

<170> PatentIn Ver. 2.1

<210> 1
<211> 2367
<212> DNA
<213> Gluconobacter melanogenus N44-1

<400> 1
atgaacagcg gcccccgcac gctctccatg atcatcggga ttctgggcgc cctcatggcc 60
gccttcctga tcatcgaagg cctccacctc atcatcctcg gcggctcgtg gttctacacc 120
ctcgccggca tcgcgctggc ggccagcagc gtctacatga tccgtcgcaa catcctctcg 180
acatggatcg ccctgggcct gcttgtggca acagccctgt ggtcgctcgc cgaagtcggc 240
accagcttct ggcccagctt ctcccgcctg atcgtgttcc tgtgcgtcgc cctgatcgcg 300
actctcatgg cgccctggct cagcggcccc ggccggcgct acttcacccg ccccgtcaca 360
ggcgccacat ccggcgccct cggcgcgatc atcgtggctt tcctcgccgg catgttccgg 420
gtccacccga ccatcgcccc gcaggacacc acccacccgc aggaaaccgc gtccaccgcc 480
gactccgacc agccaggcca tgactggccc gcctatggcc gcacggcttc cggcacgcgc 540
tacgccagct tcacgcagat caaccgcgac aatgtcagca agctccgcgt cgcctggacc 600
taccgcaccg gcgacatggc gctgaacggc gccgagttcc agggcacccc catcaagatc 660
ggcgacacgg tctatatctg ctcaccgcac aacatcgtct cggcccttga cccggacacc 720
ggcacggaaa agtggaagtt cgacccccac gcccagacga aagtctggca gcgctgccgc 780
ggcgtcggct actggcatga cagcacggcc acggacgcca acgcgccctg cgcctcgcgc 840
atcgtcctca ccacgatcga cgcccgcctc atcaccatcg acgcccgtac cggccaggcc 900
tgcacggatt tcggaacgaa cggcaacgtc aatctcctga ccggcctcgg cccgacagct 960
cccggctcgt actacccgac cgccgccccc ctcgtggcgg gtgacatcgt ggtcgtcggc 1020
ggccgcatcg ccgataacga gcgcaccggc gagccctccg gcgtcgtccg cggctatgat 1080
gtccgcaccg gcgcacaggt ctgggcctgg gacgccacca acccgcatcg cggcaccaca 1140
cctctggccg aaggcgagat ctaccccgcc gaaaccccca acatgtgggg caccgccagc 1200
tacgacccga aactcaacct cgtcttcttc ccgctcggca accagacccc cgatttctgg 1260
ggcggcgacc gcagcaaggc ctcagacgaa tacaacgacg ccttcgtcgc cgtggacgcc 1320
aagaccggcg acgaacgctg gcacttccgc accgccaacc acgacctcgt ggactacgat 1380
gccacggccc agcccatcct ctatgacatt ccggacggcc atggcggcac ccgcccggcg 1440
atcatcgcca tgaccaagcg cggccagatc ttcgtgctcg accgccgcga cggcaccccg 1500
atcgtccctg tggaaatgcg caaagtcccg caggacggcg caccggaaca ccagtacctc 1560
gcccccgaac agccctattc cgccctctcc atcggaacag agcgcctgaa acccagcgac 1620

```
atgtggggtg gtacgatctt cgaccagctc ctgtgccgca tccagttcgc ctcctaccgc 1680
tatgaaggcg agttcacccc cgtcaacgag aaacaggcca ccatcatcta tccgggctat 1740
tacggcggca tcaactgggg cggcggcgcc gtggatgaaa gcaccggaac gctgctggtc 1800
aacgacatcc gcatggccca gtggggcaag ttcatgaagc aggaagaagc ccgtcgcagc 1860
ggcttcaaac ccagctcgga aggcgaatat tccgaacaga aaggcacccc ctggggcgtc 1920
gtccgctcga tgttcttctc ccccgccggt ctcccctgcg tgaaaccgcc ctatggcacg 1980
atgaacgcca tcgacctgcg cagcggcaag gtcaaatgga gcatgccgct tggcacgatc 2040
caggacatgc cggtccacgg catggtccca ggcctcgcca tcccgctcgg aatgccgacc 2100
atgagcggcc cgctggccac ccataccggc ctggtgttct tctccggcac gctcgacaac 2160
tatgtccgcg cgctcaacac cgacaccggc gaagtcgtct ggaaagcccg tctccccgtc 2220
gcctcacagg ccgctccgat gagctacatg tccgacaaga ccggcaaaca gtacatcgtc 2280
gtcaccgcag gcggcctgac ccgctccggc gtcgacaaaa accgcggcga ctacgtcatc 2340
gcctacgccc tgccctccga agaataa                                      2367
```

<210> 2
<211> 788
<212> PRT
<213> Gluconobacter melanogenus N44-1

<400> 2

```
Met Asn Ser Gly Pro Arg Thr Leu Ser Met Ile Ile Gly Ile Leu Gly
  1               5                  10                  15

Ala Leu Met Ala Ala Phe Leu Ile Ile Glu Gly Leu His Leu Ile Ile
             20                  25                  30

Leu Gly Gly Ser Trp Phe Tyr Thr Leu Ala Gly Ile Ala Leu Ala Ala
         35                  40                  45

Ser Ser Val Tyr Met Ile Arg Arg Asn Ile Leu Ser Thr Trp Ile Ala
     50                  55                  60

Leu Gly Leu Leu Val Ala Thr Ala Leu Trp Ser Leu Ala Glu Val Gly
 65                  70                  75                  80

Thr Ser Phe Trp Pro Ser Phe Ser Arg Leu Ile Val Phe Leu Cys Val
                 85                  90                  95

Ala Leu Ile Ala Thr Leu Met Ala Pro Trp Leu Ser Gly Pro Gly Arg
            100                 105                 110

Arg Tyr Phe Thr Arg Pro Val Thr Gly Ala Thr Ser Gly Ala Leu Gly
        115                 120                 125

Ala Ile Ile Val Ala Phe Leu Ala Gly Met Phe Arg Val His Pro Thr
    130                 135                 140
```

Ile Ala Pro Gln Asp Thr Thr His Pro Gln Glu Thr Ala Ser Thr Ala
145 150 155 160

Asp Ser Asp Gln Pro Gly His Asp Trp Pro Ala Tyr Gly Arg Thr Ala
165 170 175

Ser Gly Thr Arg Tyr Ala Ser Phe Thr Gln Ile Asn Arg Asp Asn Val
180 185 190

Ser Lys Leu Arg Val Ala Trp Thr Tyr Arg Thr Gly Asp Met Ala Leu
195 200 205

Asn Gly Ala Glu Phe Gln Gly Thr Pro Ile Lys Ile Gly Asp Thr Val
210 215 220

Tyr Ile Cys Ser Pro His Asn Ile Val Ser Ala Leu Asp Pro Asp Thr
225 230 235 240

Gly Thr Glu Lys Trp Lys Phe Asp Pro His Ala Gln Thr Lys Val Trp
245 250 255

Gln Arg Cys Arg Gly Val Gly Tyr Trp His Asp Ser Thr Ala Thr Asp
260 265 270

Ala Asn Ala Pro Cys Ala Ser Arg Ile Val Leu Thr Thr Ile Asp Ala
275 280 285

Arg Leu Ile Thr Ile Asp Ala Arg Thr Gly Gln Ala Cys Thr Asp Phe
290 295 300

Gly Thr Asn Gly Asn Val Asn Leu Leu Thr Gly Leu Gly Pro Thr Ala
305 310 315 320

Pro Gly Ser Tyr Tyr Pro Thr Ala Ala Pro Leu Val Ala Gly Asp Ile
325 330 335

Val Val Val Gly Gly Arg Ile Ala Asp Asn Glu Arg Thr Gly Glu Pro
340 345 350

Ser Gly Val Val Arg Gly Tyr Asp Val Arg Thr Gly Ala Gln Val Trp
355 360 365

Ala Trp Asp Ala Thr Asn Pro His Arg Gly Thr Thr Pro Leu Ala Glu
370 375 380

Gly Glu Ile Tyr Pro Ala Glu Thr Pro Asn Met Trp Gly Thr Ala Ser
385 390 395 400

Tyr Asp Pro Lys Leu Asn Leu Val Phe Phe Pro Leu Gly Asn Gln Thr
405 410 415

Pro Asp Phe Trp Gly Gly Asp Arg Ser Lys Ala Ser Asp Glu Tyr Asn
420 425 430

Asp Ala Phe Val Ala Val Asp Ala Lys Thr Gly Asp Glu Arg Trp His
435 440 445

Phe Arg Thr Ala Asn His Asp Leu Val Asp Tyr Asp Ala Thr Ala Gln
450 455 460

Pro Ile Leu Tyr Asp Ile Pro Asp Gly His Gly Gly Thr Arg Pro Ala
465 470 475 480

Ile Ile Ala Met Thr Lys Arg Gly Gln Ile Phe Val Leu Asp Arg Arg
485 490 495

Asp Gly Thr Pro Ile Val Pro Val Glu Met Arg Lys Val Pro Gln Asp
500 505 510

Gly Ala Pro Glu His Gln Tyr Leu Ala Pro Glu Gln Pro Tyr Ser Ala
515 520 525

Leu Ser Ile Gly Thr Glu Arg Leu Lys Pro Ser Asp Met Trp Gly Gly
530 535 540

Thr Ile Phe Asp Gln Leu Leu Cys Arg Ile Gln Phe Ala Ser Tyr Arg
545 550 555 560

Tyr Glu Gly Glu Phe Thr Pro Val Asn Glu Lys Gln Ala Thr Ile Ile
565 570 575

Tyr Pro Gly Tyr Tyr Gly Gly Ile Asn Trp Gly Gly Gly Ala Val Asp
580 585 590

Glu Ser Thr Gly Thr Leu Leu Val Asn Asp Ile Arg Met Ala Gln Trp
595 600 605

Gly Lys Phe Met Lys Gln Glu Glu Ala Arg Arg Ser Gly Phe Lys Pro
610 615 620

Ser Ser Glu Gly Glu Tyr Ser Glu Gln Lys Gly Thr Pro Trp Gly Val
625 630 635 640

Val Arg Ser Met Phe Phe Ser Pro Ala Gly Leu Pro Cys Val Lys Pro
645 650 655

Pro Tyr Gly Thr Met Asn Ala Ile Asp Leu Arg Ser Gly Lys Val Lys
660 665 670

Trp Ser Met Pro Leu Gly Thr Ile Gln Asp Met Pro Val His Gly Met
675 680 685

Val Pro Gly Leu Ala Ile Pro Leu Gly Met Pro Thr Met Ser Gly Pro
690 695 700

Leu Ala Thr His Thr Gly Leu Val Phe Phe Ser Gly Thr Leu Asp Asn
705 710 715 720

Tyr Val Arg Ala Leu Asn Thr Asp Thr Gly Glu Val Val Trp Lys Ala
725 730 735

Arg Leu Pro Val Ala Ser Gln Ala Ala Pro Met Ser Tyr Met Ser Asp
740 745 750

Lys Thr Gly Lys Gln Tyr Ile Val Val Thr Ala Gly Gly Leu Thr Arg
755 760 765

Ser Gly Val Asp Lys Asn Arg Gly Asp Tyr Val Ile Ala Tyr Ala Leu
770 775 780

Pro Ser Glu Glu
785

<210> 3
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 3
cgccttctat gaaaggttgg 20

<210> 4
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 4
agcggatgga gatcgggcgg 20

<210> 5
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 5
atgaacagcg gcccccgcac gctctccatg 30

<210> 6
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 6
ccggaacatg ccggcgagga aagccacgat 30

<210> 7
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 7
tgactggccc gcctatggcc gcacggcttc 30

<210> 8
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 8
ttcttcggag ggcagggcgt aggcgatgac 30

<210> 9
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 9
cgggactttg cgcatttcca cagggacgat 30

<210> 10
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 10
agcccatcct ctatgacatt ccggacggcc 30

<210> 11
<211> 771
<212> DNA
<213> Gluconobacter melanogenus IFO 3292

<400> 11
ccgcccggcg atcatcgcca tgaccaagcg cggccagatc ttcgtgctcg accgccgcga 60
cggcaccccg atcgtccccg tggaaatgcg caaagtcccc caggacggcg caccggaaca 120
ccagtacctc gcccccgaac agccctattc cgccctctcc atcggaacag agcgcctgaa 180
acccagcgat atgtggggcg gcacgatctt cgaccagctc ctgtgccgca tccagttcgc 240
ctcctaccgc tatgaaggcg agttcacccc cgtcaacgag aagcaggcca ccatcatcta 300
tccgggctat tacggcggca tcaactgggg cggcggcgcc gtggatgaaa gcaccggaac 360
gctgctggtc aacgacatcc gcatggccca gtggggcaag ttcatgaagc aagaagaagc 420
ccgccgcagc ggcttcaaac ccagctcgga aggcgaatat tccgaacaga aaggcacccc 480
ctggggcgtc gtccgctcga tgttcttctc ccccgccggt ctcccctgcg tgaaaccgcc 540
ctatggcacg atgaacgcca tcgacctgcg cagcggcaag gtcaaatgga gcatgccgct 600
tggcacgatc caggacatgc cggtccacgg catggtcccc ggcctcgcca tcccgctcgg 660
aatgccgacc atgagcggcc cgctggccac ccataccggc ctggtcttct tctccggcac 720
gctcgacaac tatgtccgcg cgctcaacac cgacaccggc gaagtcgtct g 771

<210> 12
<211> 256
<212> PRT
<213> Gluconobacter melanogenus IFO 3292

<400> 12

```
Arg Pro Ala Ile Ile Ala Met Thr Lys Arg Gly Gln Ile Phe Val Leu
 1               5                  10                  15

Asp Arg Arg Asp Gly Thr Pro Ile Val Pro Val Glu Met Arg Lys Val
            20                  25                  30

Pro Gln Asp Gly Ala Pro Glu His Gln Tyr Leu Ala Pro Glu Gln Pro
        35                  40                  45

Tyr Ser Ala Leu Ser Ile Gly Thr Glu Arg Leu Lys Pro Ser Asp Met
    50                  55                  60

Trp Gly Gly Thr Ile Phe Asp Gln Leu Leu Cys Arg Ile Gln Phe Ala
65                  70                  75                  80

Ser Tyr Arg Tyr Glu Gly Glu Phe Thr Pro Val Asn Glu Lys Gln Ala
                85                  90                  95

Thr Ile Ile Tyr Pro Gly Tyr Tyr Gly Gly Ile Asn Trp Gly Gly Gly
            100                 105                 110

Ala Val Asp Glu Ser Thr Gly Thr Leu Leu Val Asn Asp Ile Arg Met
        115                 120                 125

Ala Gln Trp Gly Lys Phe Met Lys Gln Glu Glu Ala Arg Arg Ser Gly
    130                 135                 140

Phe Lys Pro Ser Ser Glu Gly Glu Tyr Ser Glu Gln Lys Gly Thr Pro
145                 150                 155                 160

Trp Gly Val Val Arg Ser Met Phe Phe Ser Pro Ala Gly Leu Pro Cys
                165                 170                 175

Val Lys Pro Pro Tyr Gly Thr Met Asn Ala Ile Asp Leu Arg Ser Gly
            180                 185                 190

Lys Val Lys Trp Ser Met Pro Leu Gly Thr Ile Gln Asp Met Pro Val
        195                 200                 205

His Gly Met Val Pro Gly Leu Ala Ile Pro Leu Gly Met Pro Thr Met
    210                 215                 220
```

Ser Gly Pro Leu Ala Thr His Thr Gly Leu Val Phe Phe Ser Gly Thr
225 230 235 240

Leu Asp Asn Tyr Val Arg Ala Leu Asn Thr Asp Thr Gly Glu Val Val
245 250 255

<210> 13
<211> 344
<212> DNA
<213> Gluconobacter rubiginosus IFO 3244

<400> 13
atcatcggga ttctgggcgc cctcatggcc gccttcctga tcatcgaagg cctccacctc 60
atcatcctcg gcggctcgtg gttttacacc ctcgccggca tcgcgctggc ggccagcagc 120
gtctacatga tccgtcgcaa catcctctcg acatggatcg ccctcggcct gcttgtagca 180
acagccctgt ggtcgctcgc cgaagtcggc accagcttct ggcccagctt ctcccgcctg 240
atcgtgttcc tgtgcgtcgc cctgatcgcg actctcatgg cgccctggct cagcggcccc 300
ggccggcgct acttcacccg ccccgtcaca ggggccacct ccgg 344

<210> 14
<211> 114
<212> PRT
<213> Gluconobacter rubiginosus IFO 3244

<400> 14
Ile Ile Gly Ile Leu Gly Ala Leu Met Ala Ala Phe Leu Ile Ile Glu
1 5 10 15

Gly Leu His Leu Ile Ile Leu Gly Gly Ser Trp Phe Tyr Thr Leu Ala
20 25 30

Gly Ile Ala Leu Ala Ala Ser Ser Val Tyr Met Ile Arg Arg Asn Ile
35 40 45

Leu Ser Thr Trp Ile Ala Leu Gly Leu Leu Val Ala Thr Ala Leu Trp
50 55 60

Ser Leu Ala Glu Val Gly Thr Ser Phe Trp Pro Ser Phe Ser Arg Leu
65 70 75 80

Ile Val Phe Leu Cys Val Ala Leu Ile Ala Thr Leu Met Ala Pro Trp
85 90 95

Leu Ser Gly Pro Gly Arg Arg Tyr Phe Thr Arg Pro Val Thr Gly Ala
100 105 110

Thr Ser

<210> 15
<211> 110
<212> DNA
<213> Gluconobacter rubiginosus IFO 3244

<400> 15
agcccatcct ctatgacatt ccggacggcc atggcggcac ccgcccggcg atcatcgcca 60
tgaccaagcg cggccagatc ttcgtgctcg accgccgcga cggcaccccg 110

<210> 16
<211> 36
<212> PRT
<213> Gluconobacter rubiginosus IFO 3244

<400> 16
Pro Ile Leu Tyr Asp Ile Pro Asp Gly His Gly Gly Thr Arg Pro Ala
1 5 10 15

Ile Ile Ala Met Thr Lys Arg Gly Gln Ile Phe Val Leu Asp Arg Arg
20 25 30

Asp Gly Thr Pro
35

<210> 17
<211> 801
<212> DNA
<213> Gluconobacter rubiginosus IFO 3244

<400> 17
gcgcggccag atcttcgtgc tcgaccgccg cgacggcacc ccgatcgtcc ccgtggaaat 60
gcgcaaagtc ccccaggacg gcgcaccgga acaccagtac ctcgcccccg aacagcccta 120
ttccgccctc tccatcggaa cagagcgcct gaaacccagc gatatgtggg gcggcacgat 180
cttcgaccag ctcctgtgcc gcatccagtt cgcctcctac cgctatgaag gcgagttcac 240
ccccgtcaac gagaagcagg ccaccatcat ctatccgggc tattacggcg gcatcaactg 300
gggcggcggc gccgtggatg aaagcaccgg aacgctgctg gtcaacgaca tccgcatggc 360
ccagtggggc aagttcatga agcaagaaga agcccgccgc agcggcttca aacccagctc 420
ggaaggcgaa tattccgaac agaaaggcac cccctggggc gtcgtccgct cgatgttctt 480

ctcccccgcc ggtctcccct gcgtgaaacc gccctatggc acgatgaacg ccatcgacct 540
gcgcagcggc aaggtcaaat ggagcatgcc gcttggcacg atccaggaca tgccggtcca 600
cggcatggtc cccggcctcg ccatcccgct cggaatgccg accatgagcg gcccgctggc 660
cacccatacc ggcctggtct tcttctccgg cacgctcgac aactatgtcc gcgcgctcaa 720
caccgacacc ggcgaagtcg tctggaaagc ccgtctcccc gtcgcctcac aggccgctcc 780
gatgagctac atgtccgaca a 801

<210> 18
<211> 266
<212> PRT
<213> Gluconobacter rubiginosus IFO 3244

<400> 18
Arg Gly Gln Ile Phe Val Leu Asp Arg Arg Asp Gly Thr Pro Ile Val
1 5 10 15

Pro Val Glu Met Arg Lys Val Pro Gln Asp Gly Ala Pro Glu His Gln
20 25 30

Tyr Leu Ala Pro Glu Gln Pro Tyr Ser Ala Leu Ser Ile Gly Thr Glu
35 40 45

Arg Leu Lys Pro Ser Asp Met Trp Gly Gly Thr Ile Phe Asp Gln Leu
50 55 60

Leu Cys Arg Ile Gln Phe Ala Ser Tyr Arg Tyr Glu Gly Glu Phe Thr
65 70 75 80

Pro Val Asn Glu Lys Gln Ala Thr Ile Ile Tyr Pro Gly Tyr Tyr Gly
85 90 95

Gly Ile Asn Trp Gly Gly Gly Ala Val Asp Glu Ser Thr Gly Thr Leu
100 105 110

Leu Val Asn Asp Ile Arg Met Ala Gln Trp Gly Lys Phe Met Lys Gln
115 120 125

Glu Glu Ala Arg Arg Ser Gly Phe Lys Pro Ser Ser Glu Gly Glu Tyr
130 135 140

Ser Glu Gln Lys Gly Thr Pro Trp Gly Val Val Arg Ser Met Phe Phe
145 150 155 160

Ser Pro Ala Gly Leu Pro Cys Val Lys Pro Pro Tyr Gly Thr Met Asn
165 170 175

Ala Ile Asp Leu Arg Ser Gly Lys Val Lys Trp Ser Met Pro Leu Gly

180 185 190

Thr Ile Gln Asp Met Pro Val His Gly Met Val Pro Gly Leu Ala Ile
195 200 205

Pro Leu Gly Met Pro Thr Met Ser Gly Pro Leu Ala Thr His Thr Gly
210 215 220

Leu Val Phe Phe Ser Gly Thr Leu Asp Asn Tyr Val Arg Ala Leu Asn
225 230 235 240

Thr Asp Thr Gly Glu Val Val Trp Lys Ala Arg Leu Pro Val Ala Ser
245 250 255

Gln Ala Ala Pro Met Ser Tyr Met Ser Asp
260 265

<210> 19
<211> 350
<212> DNA
<213> Gluconobacter oxydans IFO 3287

<220>
<221> misc_feature
<222> (123)
<223> n is a or c or g or t

<400> 19
atcatcggga ttctgggcgc cctcatggcc gccttcctga tcatcgaagg cctccacctc 60
atcatcctcg gcggctcatg gttttacacc ctcgccggca tcgcgctggc agccagcagc 120
gtntacatga tccgtcgcaa catcctctcg acatggatcg ccctcggcct gcttgtggca 180
acagccctgt ggtcgctcgc cgaagtcggc accagcttct ggcccagctt ctcccgcctg 240
atcgtatttc tgtgcgtcgc cctgatcgcg accctcatgg cgccctggct cagcggcccc 300
ggccggcgct acttcacccg ccccgtcaca ggcgccacct ccggcgccct 350

<210> 20
<211> 116
<212> PRT
<213> Gluconobacter oxydans IFO 3287

<400> 20
Ile Ile Gly Ile Leu Gly Ala Leu Met Ala Ala Phe Leu Ile Ile Glu
1 5 10 15

Gly Leu His Leu Ile Ile Leu Gly Gly Ser Trp Phe Tyr Thr Leu Ala
20 25 30

Gly Ile Ala Leu Ala Ala Ser Ser Val Tyr Met Ile Arg Arg Asn Ile
35 40 45

Leu Ser Thr Trp Ile Ala Leu Gly Leu Leu Val Ala Thr Ala Leu Trp
50 55 60

Ser Leu Ala Glu Val Gly Thr Ser Phe Trp Pro Ser Phe Ser Arg Leu
65 70 75 80

Ile Val Phe Leu Cys Val Ala Leu Ile Ala Thr Leu Met Ala Pro Trp
85 90 95

Leu Ser Gly Pro Gly Arg Arg Tyr Phe Thr Arg Pro Val Thr Gly Ala
100 105 110

Thr Ser Gly Ala
115

<210> 21
<211> 808
<212> DNA
<213> Gluconobacter oxydans IFO 3287

<400> 21
gcaagctccg cgtcgcctgg acctaccgca ctggcgacat ggcgctgaac ggggccgagt 60
tccagggcac ccccatcaag atcggcgaca cggtctatat ctgctcgccg cacaacatcg 120
tctcggccct cgaccccgat accggcacgg aaaagtggaa gttcgacccc cacgcccaga 180
cgaaagtctg gcagcgctgc cgcggcgtcg gctactggca tgacagcacg gccacggacg 240
ccaacgcgcc ctgcgcctcg cgcatcgtcc tcaccacgat cgacgcccgc ctcatcacca 300
tcgacgcccg caccggccag gcctgcacgg atttcggaac gaacggcaac gtcaatctcc 360
tgaccggcct cggcccgaca gcccccggtt cctactaccc gaccgccgcc cccctcgtgg 420
ccggtgacat cgtggtcgtc ggcggccgca tcgccgataa cgagcgcacc ggcgaaccct 480
ccggcgtcgt ccgcggctat gacgtccgca ccggcgcgca ggtctgggcc tgggacgcca 540
ccaacccgca tcgcggcacc acaccgctgg ccgaaggcga gatctatccc gccgaaaccc 600
ccaacatgtg gggcaccgcc agctacgacc cgaagctcaa cctcgtcttc ttcccgctcg 660
gcaaccagac ccccgatttc tggggcggcg accgcagcaa ggcttctgat gaatacaacg 720
acgccttcgt cgccgtggac gccaagaccg gcgacgaacg ctggcacttc cgcaccgcca 780
accacgacct cgtggactac gatgccac 808

<210> 22
<211> 268
<212> PRT
<213> Gluconobacter oxydans IFO 3287

<400> 22

```
Lys Leu Arg Val Ala Trp Thr Tyr Arg Thr Gly Asp Met Ala Leu Asn
 1               5                  10                  15

Gly Ala Glu Phe Gln Gly Thr Pro Ile Lys Ile Gly Asp Thr Val Tyr
            20                  25                  30

Ile Cys Ser Pro His Asn Ile Val Ser Ala Leu Asp Pro Asp Thr Gly
        35                  40                  45

Thr Glu Lys Trp Lys Phe Asp Pro His Ala Gln Thr Lys Val Trp Gln
    50                  55                  60

Arg Cys Arg Gly Val Gly Tyr Trp His Asp Ser Thr Ala Thr Asp Ala
65                  70                  75                  80

Asn Ala Pro Cys Ala Ser Arg Ile Val Leu Thr Thr Ile Asp Ala Arg
                85                  90                  95

Leu Ile Thr Ile Asp Ala Arg Thr Gly Gln Ala Cys Thr Asp Phe Gly
            100                 105                 110

Thr Asn Gly Asn Val Asn Leu Leu Thr Gly Leu Gly Pro Thr Ala Pro
        115                 120                 125

Gly Ser Tyr Tyr Pro Thr Ala Ala Pro Leu Val Ala Gly Asp Ile Val
    130                 135                 140

Val Val Gly Gly Arg Ile Ala Asp Asn Glu Arg Thr Gly Glu Pro Ser
145                 150                 155                 160

Gly Val Val Arg Gly Tyr Asp Val Arg Thr Gly Ala Gln Val Trp Ala
                165                 170                 175

Trp Asp Ala Thr Asn Pro His Arg Gly Thr Thr Pro Leu Ala Glu Gly
            180                 185                 190

Glu Ile Tyr Pro Ala Glu Thr Pro Asn Met Trp Gly Thr Ala Ser Tyr
        195                 200                 205

Asp Pro Lys Leu Asn Leu Val Phe Phe Pro Leu Gly Asn Gln Thr Pro
    210                 215                 220

Asp Phe Trp Gly Gly Asp Arg Ser Lys Ala Ser Asp Glu Tyr Asn Asp
225                 230                 235                 240

Ala Phe Val Ala Val Asp Ala Lys Thr Gly Asp Glu Arg Trp His Phe
                245                 250                 255
```

Arg Thr Ala Asn His Asp Leu Val Asp Tyr Asp Ala
260 265

<210> 23
<211> 800
<212> DNA
<213> Gluconobacter oxydans IFO 3287

<400> 23
tcttcgtgct cgaccgccgc gacggcaccc cgatcgtccc cgtggaaatg cgcaaagtcc 60
cgcaggacgg cgcaccggaa caccagtacc tcgcccccga acagccctat tccgccctct 120
ccatcggaac agagcgcctg aaacccagcg atatgtgggg tggtacgatt ttcgaccagc 180
tcctgtgccg catccagttc gcctcctacc gctatgaagg cgagttcacc cccgtcaacg 240
agaaacaggc caccatcatc tatccgggct attacggcgg catcaactgg ggcggcggcg 300
ccgtggatga aagcaccgga acgctgctgg tcaacgacat ccgcatggcc cagtggggca 360
agttcatgaa gcaggaagaa gcccgtcgca gcggcttcaa acccagctcg gaaggcgaat 420
attccgaaca gaaaggcacc ccctggggcg tcgtccgctc gatgttcttc tcccccgccg 480
gtctcccctg cgtaaaaccg ccctatggca cgatgaacgc catcgacctg cgcagcggca 540
aggtgaaatg gagcatgccg cttggcacga tccaggacat gccggtccac ggcatggtcc 600
caggcctcgc catcccgctc ggaatgccaa ccatgagcgg cccgctggcc acccataccg 660
gcttggtctt cttctccggc acgctcgaca actacgtccg cgcgctcaac accgacaccg 720
gcgaggtcgt ctggaaagcc cgtctccccg tcgcctcaca ggccgctccg atgagctaca 780
tgtccgacaa gaccggcaaa 800

<210> 24
<211> 266
<212> PRT
<213> Gluconobacter oxydans IFO 3287

<400> 24
Phe Val Leu Asp Arg Arg Asp Gly Thr Pro Ile Val Pro Val Glu Met
1 5 10 15

Arg Lys Val Pro Gln Asp Gly Ala Pro Glu His Gln Tyr Leu Ala Pro
20 25 30

Glu Gln Pro Tyr Ser Ala Leu Ser Ile Gly Thr Glu Arg Leu Lys Pro
35 40 45

Ser Asp Met Trp Gly Gly Thr Ile Phe Asp Gln Leu Leu Cys Arg Ile
50 55 60

Gln Phe Ala Ser Tyr Arg Tyr Glu Gly Glu Phe Thr Pro Val Asn Glu
65 70 75 80

Lys Gln Ala Thr Ile Ile Tyr Pro Gly Tyr Tyr Gly Gly Ile Asn Trp

85 90 95

Gly Gly Gly Ala Val Asp Glu Ser Thr Gly Thr Leu Leu Val Asn Asp
100 105 110

Ile Arg Met Ala Gln Trp Gly Lys Phe Met Lys Gln Glu Glu Ala Arg
115 120 125

Arg Ser Gly Phe Lys Pro Ser Ser Glu Gly Glu Tyr Ser Glu Gln Lys
130 135 140

Gly Thr Pro Trp Gly Val Val Arg Ser Met Phe Phe Ser Pro Ala Gly
145 150 155 160

Leu Pro Cys Val Lys Pro Pro Tyr Gly Thr Met Asn Ala Ile Asp Leu
165 170 175

Arg Ser Gly Lys Val Lys Trp Ser Met Pro Leu Gly Thr Ile Gln Asp
180 185 190

Met Pro Val His Gly Met Val Pro Gly Leu Ala Ile Pro Leu Gly Met
195 200 205

Pro Thr Met Ser Gly Pro Leu Ala Thr His Thr Gly Leu Val Phe Phe
210 215 220

Ser Gly Thr Leu Asp Asn Tyr Val Arg Ala Leu Asn Thr Asp Thr Gly
225 230 235 240

Glu Val Val Trp Lys Ala Arg Leu Pro Val Ala Ser Gln Ala Ala Pro
245 250 255

Met Ser Tyr Met Ser Asp Lys Thr Gly Lys
260 265

<210> 25
<211> 360
<212> DNA
<213> Acetobacter sp. ATCC 15164

<220>
<221> misc_feature
<222> (123)
<223> n is a or c or g or t

<400> 25
atcatcggga ttctgggcgc cctcatggcc gccttcctga tcatcgaagg cctccacctc 60

atcatcctcg gcggctcgtg gttttacacc ctcgccggca tcgcgctggc ggccagcagc 120
gtntacatga tccgtcgcaa catcctctcg acatggatcg ccctcggcct gcttgtagca 180
acagccctgt ggtcgctcgc cgaagtcggc accagcttct ggcccagctt ctcccgcctg 240
atcgtgttcc tgtgcgtcgc cctgatcgcg actctcatgg cgccctggct cagcggcccc 300
ggccggcgct acttcacccg ccccgtcaca ggggccacct ccggcgcact cggcgccatc 360

<210> 26
<211> 120
<212> PRT
<213> Acetobacter sp. ATCC 15164

<400> 26
Ile Ile Gly Ile Leu Gly Ala Leu Met Ala Ala Phe Leu Ile Ile Glu
1 5 10 15

Gly Leu His Leu Ile Ile Leu Gly Gly Ser Trp Phe Tyr Thr Leu Ala
20 25 30

Gly Ile Ala Leu Ala Ala Ser Ser Val Tyr Met Ile Arg Arg Asn Ile
35 40 45

Leu Ser Thr Trp Ile Ala Leu Gly Leu Leu Val Ala Thr Ala Leu Trp
50 55 60

Ser Leu Ala Glu Val Gly Thr Ser Phe Trp Pro Ser Phe Ser Arg Leu
65 70 75 80

Ile Val Phe Leu Cys Val Ala Leu Ile Ala Thr Leu Met Ala Pro Trp
85 90 95

Leu Ser Gly Pro Gly Arg Arg Tyr Phe Thr Arg Pro Val Thr Gly Ala
100 105 110

Thr Ser Gly Ala Leu Gly Ala Ile
115 120

<210> 27
<211> 760
<212> DNA
<213> Acetobacter sp. ATCC 15164

<400> 27
accgcgacaa tgtcagcaag ctccgcgtcg cctggaccta ccgcaccggc gacatggcgc 60
tgaacggcgc cgaattccag ggcaccccca tcaagatcgg cgatacggtc tatatctgct 120
caccccacaa catcgtctcg gccctcgacc ccgacaccgg cacggaaaag tggaagttcg 180
acccccacgc ccagacgaaa gtctggcagc gctgccgcgg cgtcggctac tggcatgaca 240

gcacagccac ggacgccaac gcgccctgcg cctcgcgcat cgtcctcacc acgatcgacg 300
cccgcctcat caccatcgac gcccgcaccg gccaggcctg cacggatttc ggaacgaacg 360
gcaacgtcaa tctcctgacc ggcctcggcc cgacagcccc cggctcctac tacccgaccg 420
ccgcccccct cgtggcgggt gacatcgtgg tcgtcggcgg ccgcatcgcc gataacgagc 480
gcacaggcga gccttccggc gtcgtccgcg gctacgacgt ccgcaccggc gcacaggtct 540
gggcctggga cgccaccaac ccgcatcgcg gcaccacacc actggccgaa ggcgagatct 600
accccgccga aacccccaac atgtggggca ccgccagcta cgacccgaaa ctcaacctcg 660
tcttcttccc gctcggcaac cagacccccg atttctgggg cggcgaccgc agcaaggcct 720
cggatgaata caacgacgcc ttcgtcgccg tggacgccaa 760

<210> 28
<211> 252
<212> PRT
<213> Acetobacter sp. ATCC 15164

<400> 28
Arg Asp Asn Val Ser Lys Leu Arg Val Ala Trp Thr Tyr Arg Thr Gly
1 5 10 15

Asp Met Ala Leu Asn Gly Ala Glu Phe Gln Gly Thr Pro Ile Lys Ile
20 25 30

Gly Asp Thr Val Tyr Ile Cys Ser Pro His Asn Ile Val Ser Ala Leu
35 40 45

Asp Pro Asp Thr Gly Thr Glu Lys Trp Lys Phe Asp Pro His Ala Gln
50 55 60

Thr Lys Val Trp Gln Arg Cys Arg Gly Val Gly Tyr Trp His Asp Ser
65 70 75 80

Thr Ala Thr Asp Ala Asn Ala Pro Cys Ala Ser Arg Ile Val Leu Thr
85 90 95

Thr Ile Asp Ala Arg Leu Ile Thr Ile Asp Ala Arg Thr Gly Gln Ala
100 105 110

Cys Thr Asp Phe Gly Thr Asn Gly Asn Val Asn Leu Leu Thr Gly Leu
115 120 125

Gly Pro Thr Ala Pro Gly Ser Tyr Tyr Pro Thr Ala Ala Pro Leu Val
130 135 140

Ala Gly Asp Ile Val Val Val Gly Gly Arg Ile Ala Asp Asn Glu Arg
145 150 155 160

Thr Gly Glu Pro Ser Gly Val Val Arg Gly Tyr Asp Val Arg Thr Gly

165 170 175

Ala Gln Val Trp Ala Trp Asp Ala Thr Asn Pro His Arg Gly Thr Thr
180 185 190

Pro Leu Ala Glu Gly Glu Ile Tyr Pro Ala Glu Thr Pro Asn Met Trp
195 200 205

Gly Thr Ala Ser Tyr Asp Pro Lys Leu Asn Leu Val Phe Phe Pro Leu
210 215 220

Gly Asn Gln Thr Pro Asp Phe Trp Gly Gly Asp Arg Ser Lys Ala Ser
225 230 235 240

Asp Glu Tyr Asn Asp Ala Phe Val Ala Val Asp Ala
245 250

<210> 29
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 29
ccgaattcag gccgaacagc agcaggtcac 30

<210> 30
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 30
gtgcctgggt acctcggtgg aggtcatgaa 30